

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 16:33:29 ; Search time 15352 Seconds  
(without alignments)  
11407.939 Million cell updates/sec

Title: US-10-840-060-142

Perfect score: 3081

Sequence: 1 aaactaaaaattgtgttcg.....gcgaatggcagcgtagt 3081

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_ev:\*  
12: gb\_un:\*  
13: gb\_vi:\*  
14: gb\_htg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3081	100.0	3081	6	CQ573345 Sequence
2	3081	100.0	3081	6	AX777288 Sequence
3	2605.2	84.6	3516	2	AY051763 Drosophil
4	1547.8	50.2	8043	6	CQ573344 Sequence
5	1547.8	50.2	35708	14	AC017670 Drosophil
6	1547.8	50.2	168583	2	AC104510 Drosophil
7	1547.8	50.2	283075	2	AE003485 Drosophil
8	1547.8	50.2	334028	14	AC116537 Drosophil
9	1117.6	36.3	180919	14	AC007521 Drosophil
10	967.8	31.4	180919	14	AC007521 Drosophil
11	568.6	18.5	2518	2	BT001278 Drosophil
12	564.6	18.3	2190	6	CQ577533 Sequence
13	564.6	18.3	4190	6	CQ577532 Sequence
14	564.6	18.3	84419	14	AC006245 Drosophil
15	564.6	18.3	180361	2	AC099036 Drosophil
16	564.6	18.3	283700	2	AE003459 Drosophil
17	563	18.3	38036	14	AC013192 Drosophil
18	505.8	16.4	1970	6	CQ578232 Sequence

19	505.8	16.4	2104	2	DQ151904
20	505.8	16.4	3970	6	CQ578231 Sequence
21	505.8	16.4	38036	14	AC013192 Drosophil
22	505.8	16.4	84419	14	AC006245 Drosophil
23	505.8	16.4	180361	2	AC099036 Drosophil
24	505.8	16.4	181463	2	AC008352 Drosophil
25	505.8	16.4	283700	2	AE003459 Drosophil
26	477	15.5	2325	9	BC006841 Mus muscu
27	464.2	15.1	3059	9	RNU44979 Rattus norv
28	461.4	15.0	2703	9	CGU11790 Cricetulus
29	454.6	14.8	1292	6	AR263768 Sequence
30	454.6	14.8	1292	6	AR412008 Sequence
31	454.6	14.8	1420	6	AR263771 Sequence
32	454.6	14.8	1421	6	AR412011 Sequence
33	454.6	14.8	2172	6	AR263774 Sequence
34	454.6	14.8	2172	6	AR412014 Sequence
35	453	14.7	2178	8	BT006759 Homo sapi
36	453	14.7	2178	11	AY889031 Synthetic
37	453	14.7	2178	11	AY891611 Synthetic
38	453	14.7	2178	11	BT007930 Synthetic
39	453	14.7	2422	8	AB072747 Macaca fa
40	453	14.7	2815	8	BC014924 Homo sapi
41	453	14.7	2823	8	BC008764 Homo sapi
42	453	14.7	2865	6	AX281799 Sequence
43	451.4	14.7	2175	8	CR450302 Homo sapi
44	449.8	14.6	1175	6	AR263769 Sequence
45	449.8	14.6	1175	6	AR412009 Sequence

ALIGNMENTS

RESULT 1	CQ573345	3081 bp	DNA	linear	PAT 02-FEB-2004
LOCUS	Sequence 1103 from Patent WO0171042.				
DEFINITION	CQ573345				
ACCESSION	CQ573345				
VERSION	CQ573345.1	GI:41637543			
KEYWORDS	Drosophila sp.				
SOURCE	Drosophila sp.				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE 1	Venter J.C., Adams M., Li, P.W. and Myers E.W.				
AUTHORS	Detection kits, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof				
TITLE	Patent: WO 0171042-A 1103 27-SEP-2001.				
JOURNAL	PE Corporation (NY) (US)				
FEATURES	Location/Qualifiers				
source	1..3081				
	/organism="Drosophila sp."				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:7242"				

ORIGIN

Query Match	100.0%	Score 3081;	DB 6;	Length 3081;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3081;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AAACTAAAAAATGTGTTGCTGACATCTGTCGCTTGCAGAACTATTTCTAGCAGATTTT	60	
Db	1	AAACTAAAAAATGTGTTGCTGACATCTGTCGCTTGCAGAACTATTTCTAGCAGATTTT	60	
Qy	61	GTGATATTTTCGTTGTCGTCGATTAATCCGCCAGTTTTTTTTTAAATGAAAGTGT	120	
Db	61	GTGATATTTTCGTTGTCGTCGATTAATCCGCCAGTTTTTTTTTAAATGAAAGTGT	120	
Qy	121	RACACATTGTACGGTTGGAGATAGCAGAGAGCCAGCGGCTCCGCTTTTCCTT	180	
Db	121	AAACATTTGTACGGTTGGAGATAGCAGAGAGCCAGCGGCTCCGCTTTTCCTT	180	

PD 11/5/01  
10/1840060  
Seg. ID 142  
Rn 1 +WO

Qy	181	TTTGTATTTCGGTTGCCAGCGCAACGAGCAAAA	CGACAGTTGGCATTTGAATTCAGACAAAC	240
Db	181	TTTTGTATTCGGTTGCCAGCGCAACGAAAA	CGACAGTTGGCATTTGAATTCAGCACAAAC	240
Qy	241	ACACATACTAAGCGCGACCCCGCAGCAGCACACACACACACTGGGACACTCGNAAAAA	300	
Db	241	ACACATACTAAGCGCGCACCCGCGAGCAGCACACACACACACTGGGACACTCGNAAAAA	300	
Qy	301	AAAAAACAGACGCTGTCTGGCGCACTTCGACAAAGCAGTTGGGTTTCGATTTAGTTGTCAATGC	360	
Db	301	AAAAAACAGACGCTGTCTGGCGCACTTCGACAAAGCAGTTGGGTTTCGATTTAGTTGTCAATGC	360	
Qy	361	CTTGAAATTCGGTTTCGGGCTTAGTTTCCACAAGTTTATCGCTCGTCAAGAAACACGAAA	420	
Db	361	CTTGAAATTCGGTTTCGGGCTTAGTTTCCACAAGTTTATCGCTCGTCAAGAAACACGAAA	420	
Qy	421	TAAATATATTTTCGACCTTAAAAATCTGACTTAAATTTGTTTTTGTATTGTTATTTATT	480	
Db	421	TAAATATATTTTCGACCTTAAAAATCTGACTTAAATTTGTTTTTGTATTGTTATTTATT	480	
Qy	481	TAGGCACATTTTGCACACCAACACGTAGTTACTACATCTACGACTTAACGGAACTCTCTCT	540	
Db	481	TAGGCACATTTTGCACACCAACACGTAGTTACTACATCTACGACTTAACGGAACTCTCTCT	540	
Qy	541	GCAAGCAGTGAAGTTGCTGCTCCATCAAGCAGTACTCGGAGTTAAACGAGGATAAGCCGG	600	
Db	541	GCAAGCAGTGAAGTTGCTGCTCCATCAAGCAGTACTCGGAGTTAAACGAGGATAAGCCGG	600	
Qy	601	GAGAAAGAAAGAGATCGGTGGAGATAGAGATATACAGGTGGAGTCAAGAGAGAGGA	660	
Db	601	GAGAAAGAAAGAGATCGGTGGAGATAGAGATATACAGGTGGAGTCAAGAGAGAGGA	660	
Qy	661	TCATGSACTATTACGGTGGGCGAGAGCGTCAAGATCAAGCGCAGGATCGCCGCGTCC	720	
Db	661	TCATGSACTATTACGGTGGGCGAGAGCGTCAAGATCAAGCGCAGGATCGCCGCGTCC	720	
Qy	721	ACATGGCGGTGGTGGCGGTGATCAACCGATCGGGCAAGTGCATCACAGTCGAAATGGTACG	780	
Db	721	ACATGGCGGTGGTGGCGGTGATCAACCGATCGGGCAAGTGCATCACAGTCGAAATGGTACG	780	
Qy	781	AGCGCGGGGAAACGAAGGGCAAGAGGTAGAACTTGGACGCGCATACTACGCTCAATCCGG	840	
Db	781	AGCGCGGGGAAACGAAGGGCAAGAGGTAGAACTTGGACGCGCATACTACGCTCAATCCGG	840	
Qy	841	AGCTAATGCAAGATACTGTGCAACAGCAGCGCGCCCGGAGCCCAAGAAACAGCCACCG	900	
Db	841	AGCTAATGCAAGATACTGTGCAACAGCAGCGCGCCCGGAGCCCAAGAAACAGCCACCG	900	
Qy	901	CGCCGATGAACCTCTCGCGTAAATCCCAACAATCGGCTATCGGTGGCAATCTCACACGCC	960	
Db	901	CGCCGATGAACCTCTCGCGTAAATCCCAACAATCGGCTATCGGTGGCAATCTCACACGCC	960	
Qy	961	GTATGACCATGGCCGGGAAACATGCTGAAACAGATCCAGGAAAGCCAGTTCGATTCGCAATC	1020	
Db	961	GTATGACCATGGCCGGGAAACATGCTGAAACAGATCCAGGAAAGCCAGTTCGATTCGCAATC	1020	
Qy	1021	CGATTGTGACGAGCAATAGCGTGAATACAAACAGCACTCCAAACACTACGCGCGCGGAG	1080	
Db	1021	CGATTGTGACGAGCAATAGCGTGAATACAAACAGCACTCCAAACACTACGCGCGCGGAG	1080	
Qy	1081	GTGTTGGCAACCAACGTCGACGACCACTGGATTACAGCGTCCACGCTACTCGCAAGCTG	1140	
Db	1081	GTGTTGGCAACCAACGTCGACGACCACTGGATTACAGCGTCCACGCTACTCGCAAGCTG	1140	
Qy	1141	CTACCGGCGCAGCAGCAAGGATTCGCTCGCGGTGCTTAAACACATTTGCCCAATC	1200	
Db	1141	CTACCGGCGCAGCAGCAAGGATTCGCTCGCGGTGCTTAAACACATTTGCCCAATC	1200	
Qy	1201	CCAGCGCGGAGCCAGTGTGCTCCGCGGCGCACAAAGGAGTCCGCACCTGCGGCGCACACCC	1260	
Db	1201	CCAGCGCGGAGCCAGTGTGCTCCGCGGCGCACAAAGGAGTCCGCACCTGCGGCGCACACCC	1260	
Qy	1261	AGGAGACTGGCGGCGCTAGTACCCGGCGATCGCAGCATTTGAAAGAGGTGGAGCGACTGA	1320	

Db	1361		AGGAGCTGGCGCGCTAGTACCGGCGATCGCA	CGCATTTGAAAGAGTGGAGCGACTGA	1320	
Qy	1321		AGGAGATCGGAGAGCGCGCCCGACAGCCCGAGATGA	AGGAGGAGAAAGGTGGCGC	1380	
Db	1321		AGGAGATCGGAGAGCGCGCCCGACAGCCCGAGATGA	AGGAGGAGAAAGGTGGCGC	1380	
Qy	1381		TGATGAACACAGGATCCGGGCAATCCAACTGGGAGAC	GGCGAAATGATACGCGAATATC	1440	
Db	1381		TGATGAACACAGGATCCGGGCAATCCAACTGGGAGAC	GGCGAAATGATACGCGAATATC	1440	
Qy	1441		AGAGCA	CGCTGGAAATTTGTCCGCTGCTCGATGGCCAGAGCGCTG	CAATCGCAAGGAGATCGATG	1500
Db	1441		AGAGCA	CGCTGGAAATTTGTCCGCTGCTCGATGGCCAGAGCGCTG	CAATCGCAAGGAGATCGATG	1500
Qy	1501		CAGTGTCCGTGGCGCAAGCGTCCCATTTAGCCGCAAG	AGAGTCAATCGCAAGGAGATCGATG	1560	
Db	1501		CAGTGTCCGTGGCGCAAGCGTCCCATTTAGCCGCAAG	AGAGTCAATCGCAAGGAGATCGATG	1560	
Qy	1561		TCATTTCCGTGGCGCAAGGACATGCTCATCTGTCAC	AGCGCGCAGCAAGTCTGACG	1620	
Db	1561		TCATTTCCGTGGCGCAAGGACATGCTCATCTGTCAC	AGCGCGCAGCAAGTCTGACG	1620	
Qy	1621		TCACCAAGTTCTTGAGAACCAAGTTTCGCTTCGACT	ACGCTTTCAACGACACGTGGC	1680	
Db	1621		TCACCAAGTTCTTGAGAACCAAGTTTCGCTTCGACT	ACGCTTTCAACGACACGTGGC	1680	
Qy	1681		ACAATGCCATGGTATACAAATACACAGCCAGCCGTT	TGTTGTAACATTTTCGAGGGCG	1740	
Db	1681		ACAATGCCATGGTATACAAATACACAGCCAGCCGTT	TGTTGTAACATTTTCGAGGGCG	1740	
Qy	1741		GAATGGCGACGTCTTCGCTACGGCCAGACGGGATCG	GGCAAAACGCAACCATCTGGCG	1800	
Db	1741		GAATGGCGACGTCTTCGCTACGGCCAGACGGGATCG	GGCAAAACGCAACCATCTGGCG	1800	
Qy	1801		GTGAGTTTAATGAAAGGTGACGACTGCAAGAACGGC	ATCTACGCCATCGCGCCAAAGG	1860	
Db	1801		GTGAGTTTAATGAAAGGTGACGACTGCAAGAACGGC	ATCTACGCCATCGCGCCAAAGG	1860	
Qy	1861		ATGCTTTGTGACCGCTGAATATCGCGGTTACCGCG	CGCATCTAGTCTCTCGGCCA	1920	
Db	1861		ATGCTTTGTGACCGCTGAATATCGCGGTTACCGCG	CGCATCTAGTCTCTCGGCCA	1920	
Qy	1921		GTTCCTTTGAGATTTACAGTGCGCAAGGTCTTCAT	CTCTGTCGACAGCAAGCAAACTGC	1980	
Db	1921		GTTCCTTTGAGATTTACAGTGCGCAAGGTCTTCAT	CTCTGTCGACAGCAAGCAAACTGC	1980	
Qy	1981		CGCTCTCGAGGATGTTAAACAGCAAGTCAGGTGGG	ACTCACCGAAGAGTGGTCTG	2040	
Db	1981		CGCTCTCGAGGATGTTAAACAGCAAGTCAGGTGGG	ACTCACCGAAGAGTGGTCTG	2040	
Qy	2041		ATGGCGTCGAGGAGTACTGAAGCTATTCAGACAGG	CGAATGCTGCGCGAATCTCGGCC	2100	
Db	2041		ATGGCGTCGAGGAGTACTGAAGCTATTCAGACAGG	CGAATGCTGCGCGAATCTCGGCC	2100	
Qy	2101		AGAGCTCGGCCAACTCCAAATCTGCTCGGTTTCG	CACGCGCTTTTCAGATTTGCTCGCGC	2160	
Db	2101		AGAGCTCGGCCAACTCCAAATCTGCTCGGTTTCG	CACGCGCTTTTCAGATTTGCTCGCGC	2160	
Qy	2161		CGCAGGCTCGACGAAGATCCATGGCAAGTTCTCG	TTTCATCGATCTGGCGGGCAATGAGC	2220	
Db	2161		CGCAGGCTCGACGAAGATCCATGGCAAGTTCTCG	TTTCATCGATCTGGCGGGCAATGAGC	2220	
Qy	2221		GGGGCTGGACACTTCTCTCGGCCGATCGGACAG	CGGTATGAGAGGTGCGAGATTAACA	2280	
Db	2221		GGGGCTGGACACTTCTCTCGGCCGATCGGACAG	CGGTATGAGAGGTGCGAGATTAACA	2280	
Qy	2281		AATCGCTGCTGGCCCTCAAGGAGTGCAATTCGT	CGTTGGCAAAACAGTGGCCCACTTCG	2340	
Db	2281		AATCGCTGCTGGCCCTCAAGGAGTGCAATTCGT	CGTTGGCAAAACAGTGGCCCACTTCG	2340	
Qy	2341		CCTTCGCTGCTCCAAACTCACCCAGGTGTCGGC	AGCTCGTTCAATGGCGAAGAGCA	2400	



```
1021 QY CGATTGTCAGCAGCAATAGCGTGAATACAAACAGCAATCCAACTACGCGCGCGGAG 1080
1021 Db |
1021 QY CGATTGTCAGCAGCAATAGCGTGAATACAAACAGCAATCCAACTACGCGCGCGGAG 1080
1021 Db |
1081 QY GTGGTGGACCAACAACGTCGACGACCACTGGATTACAGCGTCCACGCTACGCAAGCTG 1140
1081 Db |
1081 QY GTGGTGGACCAACAACGTCGACGACCACTGGATTACAGCGTCCACGCTACGCAAGCTG 1140
1081 Db |
1141 QY CTACCGGCCAGCAGCAGCAAGGATCGCTCGCGCGTGCCTTAATACACATTGCCCAATC 1200
1141 Db |
1141 QY CTACCGGCCAGCAGCAGCAAGGATCGCTCGCGCGTGCCTTAATACACATTGCCCAATC 1200
1141 Db |
1201 QY CCAGCGCGCAGCAGCTGTGTGTCGCGCGGACAAAGAGTGCCTACTGCGGCCACAAACC 1260
1201 Db |
1261 QY AGGAGCTGGCGCGCTAGTACCCGCGGATCGCAAGGAGTGCCTACTGCGGCCACAAACC 1320
1261 Db |
1261 QY AGGAGCTGGCGCGCTAGTACCCGCGGATCGCAAGGAGTGCCTACTGCGGCCACAAACC 1320
1261 Db |
1321 QY AGGAGCTGGCGCGCTAGTACCCGCGGATCGCAAGGAGTGCCTACTGCGGCCACAAACC 1380
1321 Db |
1321 QY AGGAGCTGGCGCGCTAGTACCCGCGGATCGCAAGGAGTGCCTACTGCGGCCACAAACC 1380
1321 Db |
1381 QY TGATGAACAGGATCCGCGCAATCCAACTGGGAGACGCGCAATGATACGCGAATATC 1440
1381 Db |
1381 QY TGATGAACAGGATCCGCGCAATCCAACTGGGAGACGCGCAATGATACGCGAATATC 1440
1381 Db |
1441 QY AGAGCAACGCTGGAATTTGCGCGTGTGCTCGATGGCCAGGCGCTCGATGACCATCAGATCA 1500
1441 Db |
1441 QY AGAGCAACGCTGGAATTTGCGCGTGTGCTCGATGGCCAGGCGCTCGATGACCATCAGATCA 1500
1441 Db |
1501 QY CAGTGTGCTGGCAAGCGTCCCATTAGCCGAGAGGTCATTCGCAAGGAGATCGATG 1560
1501 Db |
1501 QY CAGTGTGCTGGCAAGCGTCCCATTAGCCGAGAGGTCATTCGCAAGGAGATCGATG 1560
1501 Db |
1561 QY TCAATTTGCGTGGCGGCAAGGACATGCTCATCTGTCAGCGCGCGCAGGAGTTCGACC 1620
1561 Db |
1561 QY TCAATTTGCGTGGCGGCAAGGACATGCTCATCTGTCAGCGCGCGCAGGAGTTCGACC 1620
1561 Db |
1621 QY TCACCAAGTCTCTGGAGAACCAAGTTTCGCTTCGACTACGCTTCAACGACACGTCG 1680
1621 Db |
1681 QY ACAATGCCATGGTATACAAATACACAGCAAGCGTGGTGGTGAACCAATTTTCAGGGCG 1740
1681 Db |
1681 QY ACAATGCCATGGTATACAAATACACAGCAAGCGTGGTGGTGAACCAATTTTCAGGGCG 1740
1681 Db |
1741 QY GAATGGCGACGCTGCTTCGCTTACGCGCAGACGCGGATCGGCAAAACGACACCATGGCG 1800
1741 Db |
1741 QY GAATGGCGACGCTGCTTCGCTTACGCGCAGACGCGGATCGGCAAAACGACACCATGGCG 1800
1741 Db |
1801 QY GTGAGTTTAAATGGAAGTGCAGGATGCAAGAAACGCGCATCTACGCCATGCGGCCCAAG 1860
1801 Db |
1801 QY GTGAGTTTAAATGGAAGTGCAGGATGCAAGAAACGCGCATCTACGCCATGCGGCCCAAG 1860
1801 Db |
1861 QY ATGCTTTGTGACCTGTAATATGCGGTTCGCGGATGATGATCTAGTCTCTCGGCCA 1920
1861 Db |
1861 QY ATGCTTTGTGACCTGTAATATGCGGTTCGCGGATGATGATCTAGTCTCTCGGCCA 1920
1861 Db |
1921 QY GTTCTTTTGAGATTACAGTGGCAAGTCTTCGATCTTCGTCGACAAAGCAGAACTGC 1980
1921 Db |
1921 QY GTTCTTTTGAGATTACAGTGGCAAGTCTTCGATCTTCGTCGACAAAGCAGAACTGC 1980
1921 Db |
1981 QY CGCTCTGGAGGATGGTAAACAGCAAGTGCAGGTGGTGGGATCTACCGAGAGGTGGTGC 2040
1981 Db |
1981 QY CGCTCTGGAGGATGGTAAACAGCAAGTGCAGGTGGTGGGATCTACCGAGAGGTGGTGC 2040
1981 Db |
2041 QY ATGGCGTCGAGGAGTCTGAGCTATCCAGCAGCGCAATGCTGCCGGAATCCGGCC 2100
2041 Db |
2041 QY ATGGCGTCGAGGAGTCTGAGCTATCCAGCAGCGCAATGCTGCCGGAATCCGGCC 2100
2041 Db |
```

## RESULT 3

AY051763

LOCUS

DEFINITION

AY051763 3516 bp mRNA linear INV 27-AUG-2001  
Drosophila melanogaster LD29208 full length cDNA.

```
2101 QY AGACGTGGCGCAACTCCAAATTCGTCGCTTCGACGCGGTTTCCAGATTGTGCTCGCGC 2160
2101 Db |
2101 QY AGACGTGGCGCAACTCCAAATTCGTCGCTTCGACGCGGTTTCCAGATTGTGCTCGCGC 2160
2101 Db |
2161 QY CGCAGGGCTCGACGAAGATCCATGGCAAGTTCTCGTTTCATCGATCTGCGCGCAATGAGC 2220
2161 Db |
2161 QY CGCAGGGCTCGACGAAGATCCATGGCAAGTTCTCGTTTCATCGATCTGCGCGCAATGAGC 2220
2161 Db |
2221 QY GGGGGTGGACACATTCCTCGGCCGATCGGACAGCGGTAATGGAGGGTCCGAGATTAAACA 2280
2221 Db |
2221 QY GGGGGTGGACACATTCCTCGGCCGATCGGACAGCGGTAATGGAGGGTCCGAGATTAAACA 2280
2221 Db |
2281 QY AATCGTCTGGCGCTCAAGGAGTGCATTCTGCGTTGGGCAAAACAGTCGCGCCACTTGC 2340
2281 Db |
2281 QY AATCGTCTGGCGCTCAAGGAGTGCATTCTGCGTTGGGCAAAACAGTCGCGCCACTTGC 2340
2281 Db |
2341 QY CTTTCGCTGTCTCAAACTCAACAGGTGTGCGCGACTCGTTTCATTGGCCGAGAGAGCA 2400
2341 Db |
2341 QY CTTTCGCTGTCTCAAACTCAACAGGTGTGCGCGACTCGTTTCATTGGCCGAGAGAGCA 2400
2341 Db |
2401 QY AGACGTGCATGATAGCCATGATCTCGCGGAGTTCAGTCTCTGCGAGCACAAGCTCAACA 2460
2401 Db |
2401 QY AGACGTGCATGATAGCCATGATCTCGCGGAGTTCAGTCTCTGCGAGCACAAGCTCAACA 2460
2401 Db |
2461 QY CGCTGCGCTATGCGGATCGTCAAGGAGTGTGTCGCGGAGTTCAGGATATCGTCAAGTAT 2520
2461 Db |
2461 QY CGCTGCGCTATGCGGATCGTCAAGGAGTGTGTCGCGGAGTTCAGGATATCGTCAAGTAT 2520
2461 Db |
2521 QY CTGCGCGCGACACCGAGCCCATCGAGATCAACGACGAGGAGGAGGAGGAGCTCAACA 2580
2521 Db |
2521 QY CTGCGCGCGACACCGAGCCCATCGAGATCAACGACGAGGAGGAGGAGGAGCTCAACA 2580
2521 Db |
2581 QY TGGTGCATTCGCACTCGCATCGATCGCATTCCTCAATTCGATGACCGGCGACAGCTCGA 2640
2581 Db |
2581 QY TGGTGCATTCGCACTCGCATCGATCGCATTCCTCAATTCGATGACCGGCGACAGCTCGA 2640
2581 Db |
2641 QY ATAATCAGCGTGTCCGCGCTCTCATCACTCGGGGCGGTCAATTCACAAATTAATA 2700
2641 Db |
2641 QY ATAATCAGCGTGTCCGCGCTCTCATCACTCGGGGCGGTCAATTCACAAATTAATA 2700
2641 Db |
2701 QY ACAACAAACAGAACGCGGCAACATGGAACCTGGCCATGCTGAGTTTCGCTGAGCG 2760
2701 Db |
2701 QY ACAACAAACAGAACGCGGCAACATGGAACCTGGCCATGCTGAGTTTCGCTGAGCG 2760
2701 Db |
2761 QY AACACGAGATGTCGACGAGCTGATGTCGACGACCAAGGCGCATCGACGCTGAGCGA 2820
2761 Db |
2761 QY AACACGAGATGTCGACGAGCTGATGTCGACGACCAAGGCGCATCGACGCTGAGCGA 2820
2761 Db |
2821 QY CGGAGGAGATGTTGTTGGAGTATCATCGACCGTTAATGCCACTGGAGACCTTCCTCG 2880
2821 Db |
2821 QY CGGAGGAGATGTTGTTGGAGTATCATCGACCGTTAATGCCACTGGAGACCTTCCTCG 2880
2821 Db |
2881 QY CCGAGTCCGAGGCGCTGTACAATCTGACCAACTATGTGGACTACGACGAGGACTTCGTACT 2940
2881 Db |
2881 QY CCGAGTCCGAGGCGCTGTACAATCTGACCAACTATGTGGACTACGACGAGGACTTCGTACT 2940
2881 Db |
2941 QY GCAAAACGGGCGAGTGCATGTTCTCGCAGCTGTGGAATCGCCATCCAGTTCGCGCGACA 3000
2941 Db |
2941 QY GCAAAACGGGCGAGTGCATGTTCTCGCAGCTGTGGAATCGCCATCCAGTTCGCGCGACA 3000
2941 Db |
3001 QY TGATGGCCGAATATCGCGCAAGTTGGCCAAAGGAGGAGATGCTGTCGTGCGAGCTTCAATT 3060
3001 Db |
3001 QY TGATGGCCGAATATCGCGCAAGTTGGCCAAAGGAGGAGATGCTGTCGTGCGAGCTTCAATT 3060
3001 Db |
3061 QY CGCCGAATGGCAAGCTTACT 3081
3061 Db |
3061 QY CGCCGAATGGCAAGCTTACT 3081
3061 Db |
```



ACCESSION AY051763.1 GI:15291836  
VERSION FLI\_CDNA  
KEYWORDS Drosophila melanogaster (fruit fly)  
SOURCE Drosophila melanogaster  
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 3516)  
REFERENCE Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Friese, E., George, R., Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J., Nunoo, J., Pacle, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celniker, S.  
Direct Submission  
Submitted (10-AUG-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720  
This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to [cdna@fruitfly.berkeley.edu](mailto:cdna@fruitfly.berkeley.edu).

FEATURES  
source Location/Qualifiers  
1..3516  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/strain="y; cn bw sp"  
/db\_xref="taxon:7227"  
/map="10A6-10A7"  
1..3516  
/genes="Klpl10A"  
/notes="alignment with genomic scaffold AE003485"  
/db\_xref="FLYBASE:FBgn0030268"  
337..2754  
/genes="Klpl10A"  
/notes="Longest ORF"  
/codon\_start=1  
/product="Lb29208p"  
/protein\_id="AAK93187.1"  
/db\_xref="GI:15291837"  
/db\_xref="FLYBASE:FBgn0030268"  
/translation="MDMITVGSVKIKRTDGRVHMAVAVINOSGKCIIVWYERGET KGKVELDALINPELMQDTVEHAAPFKQATAPMNLNRPQTQSAIGNLTSMRT MAGCNLTKIQESQIPNPIVSSNVNTNSNTAGGCGGTTSTTTGLRPRYSAA TCQOQTRIASAVPNTLPNSAASAPAAQGVATAATTQAGAGASTRESHAKEVER LKENEKRRAQAEKKEKVALMNQDPNPNETAGMIREYSTLEFVPLLDGQAVDD HQITCVKRPISRKEVNRKEIDVISVPRKDMLI VHEPRSKVDLT KFLNHRFREDYA FNDTCDNMYKYTAKPLVKTIFEGGMATCFPAYQTGSGKTHMTGSEFNGKVDCKNG IYAMAKQVVDGVEEVKLIOHGNAASTSGTSANSRSRSHAVFOI VLRPOGSTKIHG KFSFIDLACNREGVDTSSADROTMEGAIEKSLALKEICIRALKQSAHLPPRVSKL TQVLRDSFGEKSKTMIAWISPLSSCEHTLNTLYADRVKELVVKDIVEVCPGDDT SPIETDDEESELNNVPHSHQPNNSHAPASQNNORAPASHSHGVVHNNNNN KNGNAGMLDLAMLSISEHMSDELIVHOAIIDDLQOETEMVVEYHRTVNALETFLA ESKALYNLTNYVDYDQDSYCKRGESMFSQLDIAIQCRDMMAEYRAKLKEBMLSCSF NSPNGRK"

ORIGIN

Query Match 84.6%; Score 2605.2; DB 2; Length 3516;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 2625; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 424 AATTATTTTCGACCTAAATAATCTGACTAAATTTGTGTTTTTGTATGTTATTTAG 483  
DB 98 AGTTTTTTTTTAAATGGAAAGTGCTAAACATTTGTAGCGTTGGAAAGATAGCGAAG 157  
QY 484 GCACATTTTGCACACCAACAGTAGTTACTACATCTACGACTAACCGAACTCTCTCTGCA 543  
DB 158 GCACATTTTGCACACCAACAGTAGTTACTACATCTACGACTAACCGAACTCTCTCTGCA 217  
QY 544 AGCAGTGGAAAGTTGCTGCTCCATCAAGCAGTACTCGAGTTTAAAGCAGGATAAGCGGAG 603  
DB 218 AGCAGTGGAAAGTTGCTGCTCCATCAAGCAGTACTCGAGTTTAAAGCAGGATAAGCGGAG 277  
QY 604 AAAGAGAAAGAGATCGGTGGAGAAATAGAGATATACAGTGTGAGTCAAAAGAGGAGATCA 663  
DB 278 AAAGAGAAAGAGATCGGTGGAGAAATAGAGATATACAGTGTGAGTCAAAAGAGGAGATCA 337  
QY 664 TGGACATGATTACGGTGGGGCAGAGCGTCAAGATCAAGCGGACGGATGGCCGCTCCACA 723  
DB 338 TGGACATGATTACGGTGGGGCAGAGCGTCAAGATCAAGCGGACGGATGGCCGCTCCACA 397  
QY 724 TGGCCCTGCTGGCGGTGATCAACCACTCGGGCAAGTGCATCACAGTCGAATGTGACGAGC 783  
DB 398 TGGCCCTGCTGGCGGTGATCAACCACTCGGGCAAGTGCATCACAGTCGAATGTGACGAGC 457  
QY 784 GCGGCGAAACGAAGGCGCAAGGAGTGAAGCTGAGCGCCATCTACGCTCAATCCGGAGC 843  
DB 458 GCGGCGAAACGAAGGCGCAAGGAGTGAAGCTGAGCGCCATCTACGCTCAATCCGGAGC 517  
QY 844 TAATGCAAGATCTGTCGAACAGCACGCCGCCCGGAGCCCAAGAAACCAAGCACCAGCGC 903  
DB 518 TAATGCAAGATCTGTCGAACAGCACGCCGCCCGGAGCCCAAGAAACCAAGCACCAGCGC 577  
QY 904 CGATGAACCTCTCGCGTAAATCCACAACTCGGCTATCGGTGGCAATCTCACAGCGCTA 963  
DB 578 CGATGAACCTCTCGCGTAAATCCACAACTCGGCTATCGGTGGCAATCTCACAGCGCTA 637  
QY 964 TGACCATGCGCGGAAACATGCTGAACAAGATCAGAGAAAGCCAGTGCATTTCCCAATCCGA 1023  
DB 638 TGACCATGCGCGGAAACATGCTGAACAAGATCAGAGAAAGCCAGTGCATTTCCCAATCCGA 697  
QY 1024 TTGTACAGCAATAGCTGAATACAAACAGCAATCTCCAACTACGCGCGCGGAGGTG 1083  
DB 698 TTGTACAGCAATAGCTGAATACAAACAGCAATCTCCAACTACGCGCGCGGAGGTG 757  
QY 1084 GTGGCACCAACAGCTCGACGACCACTGGATTACAGGTTACAGGTTACAGGTTACAGGTTAC 1143  
DB 758 GTGGCACCAACAGCTCGACGACCACTGGATTACAGGTTACAGGTTACAGGTTACAGGTTAC 817  
QY 1144 CCGGCGCAGCAGCAGCAAGGATCGCTCGCGCGTCTTAATACACATTTGCCCAATCCCA 1203  
DB 818 CCGGCGCAGCAGCAGCAAGGATCGCTCGCGCGTCTTAATACACATTTGCCCAATCCCA 877  
QY 1204 GCGCGCAGCAGCTAGTGTGTCGCGGCGCAAGAGAGTGCCTACGCGCGCACAACCCAGG 1263  
DB 878 GCGCGCAGCAGCTAGTGTGTCGCGGCGCAAGAGAGTGCCTACGCGCGCACAACCCAGG 937  
QY 1264 GAGCTGCGCGCGTAGTATCCCGCGCATCGCAAGATTGAAGAGGTGGAGCGCTGAAGG 1323  
DB 938 GAGCTGCGCGCGTAGTATCCCGCGCATCGCAAGATTGAAGAGGTGGAGCGCTGAAGG 997  
QY 1324 AGAATCGGAGAGCGAGCGCGCCGACAGCGCGGAGTGAAGGAGAGAGGTGGCGCTGA 1383  
DB 998 AGAATCGGAGAGCGAGCGCGCCGACAGCGCGGAGTGAAGGAGAGAGGTGGCGCTGA 1057  
QY 1384 TGAACCAAGGATCCGGCGCAATCAAACTGGGAGACGCGCAAAATGATACGCAATATCAGA 1443  
DB 1058 TGAACCAAGGATCCGGCGCAATCAAACTGGGAGACGCGCAAAATGATACGCAATATCAGA 1117  
QY 1444 GCACGCTGGAAATTTGTGCGCTGCTCGATGGCGACGCGCTCGATGATGATACATCAGACAG 1503

Db	1118	GCAGCTTGGAAATTTGTGCGCGTCTCGATGCGCCAGGCGGTGAGACCATGATCAACAG	1177
Qy	1504	TGTCGCTGCGCAAGCGTCCCAATTAGCGCGCAAGGAGGTCAATTCGCAAGGAGATCGATGTCA	1563
Db	1178	TGTCGCTGCGCAAGCGTCCCAATTAGCGCGCAAGGAGGTCAATTCGCAAGGAGATCGATGTCA	1237
Qy	1564	TTTTCGCTGCGCGCAAGGACATGCTCATCTGTCACGAGCGCGCAGCAAGGTTCGACCTCA	1623
Db	1238	TTTTCGCTGCGCGCAAGGACATGCTCATCTGTCACGAGCGCGCAGCAAGGTTCGACCTCA	1297
Qy	1624	CCAAAGTTCCTGGAGAACCAAGTTTCGCTTCGACTACGCTTCGCACTAGCGCTTCAACGCA	1683
Db	1298	CCAAAGTTCCTGGAGAACCAAGTTTCGCTTCGACTACGCTTCGCACTAGCGCTTCAACGCA	1357
Qy	1684	ATGCGCATGGTATACAAATACACAGCCAAAGCGTTTGGTGAATAACATTTTCGAGGGCGGAA	1743
Db	1358	ATGCGCATGGTATACAAATACACAGCCAAAGCGTTTGGTGAATAACATTTTCGAGGGCGGAA	1417
Qy	1744	TGGCGACGTGCTTCGCTCCTACGCGCAGACGGGATCGGGCAAAACGCACTATGGGGCGGTG	1803
Db	1418	TGGCGACGTGCTTCGCTCCTACGCGCAGACGGGATCGGGCAAAACGCACTATGGGGCGGTG	1477
Qy	1804	AGTTTAATGAAAGGTTCAGGACTGCAAGAAACGGCATCTACGCCATTCGCGGCCAAGGATG	1863
Db	1478	AGTTTAATGAAAGGTTCAGGACTGCAAGAAACGGCATCTACGCCATTCGCGGCCAAGGATG	1537
Qy	1864	TCCTTGTGACCTCGAATATGCGCGTTTACCGCGCCATGAACTCTAGTCGTCTCGGCCAGTT	1923
Db	1538	TCCTTGTGACCTCGAATATGCGCGTTTACCGCGCCATGAACTCTAGTCGTCTCGGCCAGTT	1597
Qy	1924	TCCTTGTGAGTTTACAGTGGCAAGGTCTTCGATCTTCTGTCGCAAGACGAAACTCGCGG	1983
Db	1598	TCCTTGTGAGTTTACAGTGGCAAGGTCTTCGATCTTCTGTCGCAAGACGAAACTCGCGG	1657
Qy	1984	TCCTGAGGATGTTAAACAGCAATGTCAGTGTGGGACTCAACGAGAAAGTTCGATG	2043
Db	1658	TCCTGAGGATGTTAAACAGCAATGTCAGTGTGGGACTCAACGAGAAAGTTCGATG	1717
Qy	2044	CGGTGCGAGGAGTACTGAAGCTCATCCAGCACGGCAATGCTGCCGCAATCCGGCCAGA	2103
Db	1718	CGGTGCGAGGAGTACTGAAGCTCATCCAGCACGGCAATGCTGCCGCAATCCGGCCAGA	1777
Qy	2104	CGTCCGCCAACTCCAAATTCGTTCGCGTTTCGCAACCGCGTTTTCAGATTTGTCGCGCCCG	2163
Db	1778	CGTCCGCCAACTCCAAATTCGTTCGCGTTTCGCAACCGCGTTTTCAGATTTGTCGCGCCCG	1837
Qy	2164	AGGCTTCGACGAAATCCATGGCAAGTTCTGTTTCATTCGATCTGGCGGGCAATGACGGG	2223
Db	1838	AGGCTTCGACGAAATCCATGGCAAGTTCTGTTTCATTCGATCTGGCGGGCAATGACGGG	1897
Qy	2224	CGCTGGACACTTCCTCGCGCGATTCGCGAGACGGTATGGAGGTTCGAGATTAACAAT	2283
Db	1898	CGCTGGACACTTCCTCGCGCGATTCGCGAGACGGTATGGAGGTTCGAGATTAACAAT	1957
Qy	2284	CGCTGCTGGCCCTCAAGGAGTGCATTCGTGCGTTTGGCAAAACAGTCGGCCACTTCGCCCT	2343
Db	1958	CGCTGCTGGCCCTCAAGGAGTGCATTCGTGCGTTTGGCAAAACAGTCGGCCACTTCGCCCT	2017
Qy	2344	TCCGTGTCCTCAAACTCACCCAGGTGTCGCGACTCGTTTCATTTGGCGAGAGAGCAAGA	2403
Db	2018	TCCGTGTCCTCAAACTCACCCAGGTGTCGCGACTCGTTTCATTTGGCGAGAGAGCAAGA	2077
Qy	2404	CGTGCATGATAGCCATGATCTCCGCGGACTTAGCTCTCGCGAGCACACGCTCAACACCG	2463
Db	2078	CGTGCATGATAGCCATGATCTCCGCGGACTTAGCTCTCGCGAGCACACGCTCAACACCG	2137
Qy	2464	TGCGCTATGCGGATCGTGTCAAGGAGCTGGTGGTCAAGGATATCGTCGAAGTTTGGCCCTG	2523
Db	2138	TGCGCTATGCGGATCGTGTCAAGGAGCTGGTGGTCAAGGATATCGTCGAAGTTTGGCCCTG	2197
Qy	2524	GCGGCGACACCGAGCCCATCGAGATCACGGAACGAGGAGGAGGAGCTCAACATGG	2583

Db	2198	GC	CGCGCA	CCGAGCCCATCGAGTACCGGACGAGGAGGAGGAGTCAACATGG	2255
Qy	2584	TG	CATCCGCACTCGCATCAGCTGCATCCCAATTCGATGCA	CGCGCGCAGCGAGTCTGCAATA	2643
Db	2258	TG	CATCCGCACTCGCATCAGCTGCATCCCAATTCGATGCA	CGCGCGCAGCGAGTCTGCAATA	2317
Qy	2644	AT	CAGCGTGTCTCGGCGCTCTCATCAGCTCGGGGCGGTCTATTC	CAACAATAATAAACA	2703
Db	2318	AT	CAGCGTGTCTCGGCGCTCTCATCAGCTCGGGGCGGTCTATTC	CAACAATAATAAACA	2377
Qy	2704	ACA	ACAAGACGGAACCGCGGCAACATGGA	CTGCGCCATGCTGAGTTCGCTGAGCGAAC	2763
Db	2378	ACA	ACAAGACGGAACCGCGGCAACATGGA	CTGCGCCATGCTGAGTTCGCTGAGCGAAC	2437
Qy	2764	AC	GAGATGTCGACGAGCTGATGTGTCAGCACACAGCGCATTCG	ACGAGCTTCGACGACGCG	2823
Db	2438	AC	GAGATGTCGACGAGCTGATGTGTCAGCACACAGCGCATTCG	ACGAGCTTCGACGACGCG	2497
Qy	2824	AG	GAGATGTTGGTGGAGTATCATCGCACCGCTTAATGCCACAT	TGGAGACTTCTCTCGCG	2883
Db	2498	AG	GAGATGTTGGTGGAGTATCATCGCACCGCTTAATGCCACAT	TGGAGACTTCTCTCGCG	2557
Qy	2884	AGT	CGAAGGCGCTGTACAACTGACCAACATATGTGAGTA	ACGACAGGACTCGTACTGCA	2943
Db	2558	AGT	CGAAGGCGCTGTACAACTGACCAACATATGTGAGTA	ACGACAGGACTCGTACTGCA	2617
Qy	2944	AAC	GGGCGAGTCGATGTTCTTCGACGCTCTGAGCATGCCATCC	AGTCCAGTCCGCGACATGA	3003
Db	2618	AAC	GGGCGAGTCGATGTTCTTCGACGCTCTGAGCATGCCATCC	AGTCCAGTCCGCGACATGA	2677
Qy	3004	TGG	CCGAATATCGCGCAAGTTGGCCAAAGGAGAGATGCTGT	CGTGCAGCTTCAATTCG	3063
Db	2678	TGG	CCGAATATCGCGCAAGTTGGCCAAAGGAGAGATGCTGT	CGTGCAGCTTCAATTCG	2737
Qy	3064	CGA	ATGCAAGCGTTAGT	3081	
Db	2738	CGA	ATGCAAGCGTTAGT	2755	
RESULT 4					
LOCUS	CQ573344		8043 bp	DNA	linear
DEFINITION	Sequence	1102	from Patent	WO0171042.	
ACCESSION	CQ573344				
VERSION	CQ573344.1		GI:41637542		
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
ORIGIN					
Query Match					
Best Local Similarity					
Matches 1552; Conservative					
Qy	1556	AG	CACGGGCGCTCCGTTTTTCTTTTCTTTTATTCGTTGCGCAGACGCAACGAAACGACA	215	
Db	1599	AG	CCACGGGCGCTCCGTTTTTCTTTTCTTTTATTCGTTGCGCAGACGCAACGAAACGACA	1658	
Qy	216	GT	TGGCATTTGAATTCAGCACAAACACATACTAACGCGGACCGCGACGACACAC	275	

Db 1659 GTTGGCATTTGAATTCAGCAACAACACACATCTAAGCCGACCCGCAAGCAGCACAC 1718  
Qy 276 ACACACACTGGGACACTCGAAAAAACAACAGACGCTGCGGCGACCTCGCAACAGCAG 335  
Db 1719 ACACACACTGGGACACTCGAAAAAACAACAGACGCTGCGGCGACCTCGCAACAGCAG 1778  
Qy 336 TTCGGTTCGATTTAGTTGTCATGCTTGAATTCGGTTCCGGTTCGGGCTTAGTTTCACAAGTT 395  
Db 1779 TTCGGTTCGATTTAGTTGTCATGCTTGAATTCGGTTCCGGTTCGGGCTTAGTTTCACAAGTT 1838  
Qy 396 TATCGCTCGTCAAGAAACAACAGAAATAAAATATATTTTCGACCTAAATAATCTCACTAAAT 455  
Db 1839 TATCGCTCGTCAAGAAACAACAGAAATAAAATATATTTTCGACCTAAATAATCTCACTAAAT 1898  
Qy 456 TGTGTTTTTTGTTATGATTTATTTAGGCACATTTTGCACACACACAGCTAGTTACTAC 515  
Db 1899 TGTGTTTTTTGTTATGATTTATTTAGGCACATTTTGCACACACACAGCTAGTTACTAC 1958  
Qy 516 ATCTAGACTAACCGAACTCTCTCGACGACGTGGAAGTTGCTGTCATCAAGCAGTAC 575  
Db 1959 ATCTAGACTAACCGAACTCTCTCGACGACGTGGAAGTTGCTGTCATCAAGCAGTAC 2018  
Qy 576 TCGGAGTTACGCAAGTAAGCCGCGAGAGAAAGAGATCGGTGGAGATAGAGATA 635  
Db 2019 TCGGAGTTACGCAAGTAAGCCGCGAGAGAAAGAGATCGGTGGAGATAGAGATA 2078  
Qy 636 TACAGTGGAGTCAAGAGAGAGAGATCATGGACATGATTACGGTGGGCGAGCGTCAAG 695  
Db 2079 TACAGTGGAGTCAAGAGAGAGAGATCATGGACATGATTACGGTGGGCGAGCGTCAAG 2138  
Qy 696 ATCAAGCGGACGAGTGGCGGCTGCACATGCGCGTGGTGGCGGTGATCAACAGTGGGC 755  
Db 2139 ATCAAGCGGACGAGTGGCGGCTGCACATGCGCGTGGTGGCGGTGATCAACAGTGGGC 2198  
Qy 756 AAGTGCATCACAGTCAATCGGAGCTAAATGCAAGATACTGTGCAACAGCAGCGCGCC 815  
Db 2199 AAGTGCATCACAGTCAATCGGAGCTAAATGCAAGATACTGTGCAACAGCAGCGCGCC 2258  
Qy 816 GAGCGCATACTCACGCTCAATCGGAGCTAAATGCAAGATACTGTGCAACAGCAGCGCGCC 875  
Db 2259 GAGCGCATACTCACGCTCAATCGGAGCTAAATGCAAGATACTGTGCAACAGCAGCGCGCC 2318  
Qy 876 CCGGAGCCCAAGAAACAAGCAGCCGCGCGATGAACCTCTCGGTAATCCCAACAATCG 935  
Db 2319 CCGGAGCCCAAGAAACAAGCAGCCGCGCGATGAACCTCTCGGTAATCCCAACAATCG 2378  
Qy 936 GCTATCGGTGGCAATCTCACAGCGTATGACCATGGCGGAAACATGCTGAACAGATC 995  
Db 2379 GCTATCGGTGGCAATCTCACAGCGTATGACCATGGCGGAAACATGCTGAACAGATC 2438  
Qy 996 CAGGAAAGCCAGTTCGATTCCTCAATCCGATTTGTGACGAGCAATAGCGTGAATCAACACAGC 1055  
Db 2439 CAGGAAAGCCAGTTCGATTCCTCAATCCGATTTGTGACGAGCAATAGCGTGAATCAACACAGC 2498  
Qy 1056 AACTCCAAACACTACGCGCGGAGGTGGTGACCAACAGCTCGACGACCACTGGATTA 1115  
Db 2499 AACTCCAAACACTACGCGCGGAGGTGGTGACCAACAGCTCGACGACCACTGGATTA 2558  
Qy 1116 CAGCGTCCAGGTACTCGCAAGTGTACCGGCGGACGACGACAGGATCGCTCGGG 1175  
Db 2559 CAGCGTCCAGGTACTCGCAAGTGTACCGGCGGACGACGACAGGATCGCTCGGG 2618  
Qy 1176 GTGCTTAATAACACATTCGATTCCTCAATCCGAGCGGACGAGTGTGTTCCGCGGACAA 1235  
Db 2619 GTGCTTAATAACACATTCGATTCCTCAATCCGAGCGGACGAGTGTGTTCCGCGGACAA 2678  
Qy 1236 GAGTGCCTCCTGCGGCAAAACCCAGGAGTGGCGGCTAGTACCCGCGCATGCGAC 1295  
Db 2679 GAGTGCCTCCTGCGGCAAAACCCAGGAGTGGCGGCTAGTACCCGCGCATGCGAC 2738  
Qy 1296 GCATTTGAAGAGGTGGAGCACTGAAGAGATTCGGAGAGCGGACGCGCCGACAGGCC 1355

Db 2739 GCATTTGAAGAGGTGGAGCACTGAAGAGGAATTCGCGAAGACGACGCGCCGACAGGCC 2798  
Qy 1356 GAGATGAAGAGAGAGAGGTGGCGCTGATGAACACAGGATCCGGGCAATCCAACTCGGAG 1415  
Db 2799 GAGATGAAGAGAGAGAGGTGGCGCTGATGAACACAGGATCCGGGCAATCCAACTCGGAG 2858  
Qy 1416 ACGGCGCAATGATACGCGAATATCAGAGACGCTGGAAATTTGTGCGGCTGCTCGATGGC 1475  
Db 2859 ACGGCGCAATGATACGCGAATATCAGAGACGCTGGAAATTTGTGCGGCTGCTCGATGGC 2918  
Qy 1476 CAGGCGCTGATGACCATCAGATCAGATGTCGTCGCAAGCTCCCATTAGCCGCAAG 1535  
Db 2919 CAGGCGCTGATGACCATCAGATCAGATGTCGTCGCAAGCTCCCATTAGCCGCAAG 2978  
Qy 1536 GAGTCAATTCGCAAGAGATCGATGTCATTTTCGGTCCGCGCAAGCATGCTCATCGTG 1595  
Db 2979 GAGTCAATTCGCAAGAGATCGATGTCATTTTCGGTCCGCGCAAGCATGCTCATCGTG 3038  
Qy 1596 CAGGCGCGCGGACGAGGTGACCTCACCAGTTCTCGAGAACCAACAAGTTTCGCTTC 1655  
Db 3039 CAGGCGCGCGGACGAGGTGACCTCACCAGTTCTCGAGAACCAACAAGTTTCGCTTC 3098  
Qy 1656 GACTAGCCTTCAACGACACGTCGACATGTCATGTCATGATACAAATACACAGCCAGCC 1714  
Db 3099 GACTAGCCTTCAACGACACGTCGACATGTCATGTCATGATACAAATGAGTATCCAAATCC 3157

RESULT 5  
AC017670  
LOCUS  
DEFINITION  
AC017670  
AC017670.1 GI:6554327  
VERSION  
HTG; HTGS, PHASE2.  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster (fruit fly)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 35708)  
Adams, M. and Venter, J.C.  
Direct Submission  
Submitted (10-DEC-1999) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
This sequence was identified as CDM:10211672 by the submitter.  
For more information on this record e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* this sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
1..35708  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"

ORIGIN  
Query Match 50.2%; Score 1547.8; DB 14; Length 35708;  
Best Local Similarity 99.6%; Pred. No. 2.5e-286;  
Matches 1552; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 156 AGCCAGCGGCTGCGGTTTTTCTTTTATCGTTGCCAGCGCAACGAAACGACA 215  
Db 2212 AGCCAGCGGCTGCGGTTTTTCTTTTATCGTTGCCAGCGCAACGAAACGACA 2271  
Qy 216 GTTGGCTTTGATTCAGCACAAACACATCTAACCGCGACCCGCGAGCAGCACAC 275  
Db 2272 GTTGGCTTTGATTCAGCACAAACACATCTAACCGCGACCCGCGAGCAGCACAC 2331  
Qy 276 ACACACACTGGGACACTCGAAAAAACAACAGACGCTGTCGGGAGCTCGACAGGAG 335  
Db 2332 ACACACACTGGGACACTCGAAAAAACAACAGACGCTGTCGGGAGCTCGACAGGAG 2391

336 TTGGGTTCCGATTAGTTGTCATGCTTGAATTCGGTTCCGGGCTTAGTTTCCCAAGTT 395  
3392 TTGGGTTCCGATTAGTTGTCATGCTTGAATTCGGTTCCGGGCTTAGTTTCCCAAGTT 2451  
396 TATCGCTCGTCAAGAAAACAAGAAATAAATAATTTATTTTGCACCTAAATAATCTGACTAAAT 455  
2452 TATCGCTCGTCAAGAAAACAAGAAATAAATAATTTATTTTGCACCTAAATAATCTGACTAAAT 2511  
456 TGTGTTTTTTTGTATGTTATTTATTTAGGCACATTTTGCACACCAACCAAGTAGTTACTAC 515  
2512 TGTGTTTTTTTGTATGTTATTTATTTAGGCACATTTTGCACACCAACCAAGTAGTTACTAC 2571  
516 ATCTACGACTAAGCGAACTCTCTCTGCAAGCAGTGGAAAGTTGCTGTCATCAAGCAGTAC 575  
2572 ATCTACGACTAAGCGAACTCTCTCTGCAAGCAGTGGAAAGTTGCTGTCATCAAGCAGTAC 2631  
576 TCGGAGTTAAACGACGATAAGCCGGGAGAAAGAGAAAGATCGGTGAGAAATAGAGATA 635  
2632 TCGGAGTTAAACGACGATAAGCCGGGAGAAAGAGAAAGATCGGTGAGAAATAGAGATA 2691  
636 TACAGGTGAGTCAAGAGAGAGATCATGGACATGATTACGGTGGGCGCAGAGCTCAAG 695  
2692 TACAGGTGAGTCAAGAGAGAGATCATGGACATGATTACGGTGGGCGCAGAGCTCAAG 2751  
696 ATCAAGCGGACGGATGGCGCGTCCACATGCGGTGGTGGCGGTGATCAACCAAGTCCGGC 755  
2752 ATCAAGCGGACGGATGGCGCGTCCACATGCGGTGGTGGCGGTGATCAACCAAGTCCGGC 2811  
756 AAGTGCATCAGATCGAATGTTACAGCGCGCGGAAACAGAGGGGAGGAGGTAGAACTG 815  
2812 AAGTGCATCAGATCGAATGTTACAGCGCGCGGAAACAGAGGGGAGGAGGTAGAACTG 2871  
816 GACGCGTACTCAGCTCAATCCGAGCTAATGCAAGATCTGTGCAACAGCAGCGCGCC 875  
2872 GACGCGTACTCAGCTCAATCCGAGCTAATGCAAGATCTGTGCAACAGCAGCGCGCC 2931  
876 CGGAGGCCAAGAAACAAGCCACCGCGCGATGAACCTCTCGCTGAATCCCAACAATCG 935  
2932 CGGAGGCCAAGAAACAAGCCACCGCGCGATGAACCTCTCGCTGAATCCCAACAATCG 2991  
936 GCTATCGGTGGATCTCAGACGCGGTATGACCATGGCGGAAACATGCTGAACAAGATC 995  
2992 GCTATCGGTGGATCTCAGACGCGGTATGACCATGGCGGAAACATGCTGAACAAGATC 3051  
996 CAGGAAAGCCAGTCCGATCCCAATCCGATTTGTCAGCAGCAATAGGTTGAATACAAACAGC 1055  
3052 CAGGAAAGCCAGTCCGATTTCCCAATCCGATTTGTCAGCAGCAATAGGTTGAATACAAACAGC 3111  
1056 AACTCCAACTACGCGCGCGGAGGTGGTGGCACCAACGTCAGCACCACTGGATTA 1115  
3112 AACTCCAACTACGCGCGCGGAGGTGGTGGCACCAACGTCAGCACCACTGGATTA 3171  
1116 CAGGTCACGGTACTCGAAGTCTACCGGCGCAGCAGCAGCAGATCGCTCGGCG 1175  
3172 CAGGTCACGGTACTCGAAGTCTACCGGCGCAGCAGCAGCAGATCGCTCGGCG 3231  
1176 GTGCTTAATAACATTCGCAATCCAGCGCGGAGGTGGTGGCACCAACGTCAGCACCACTGGATTA 1235  
3232 GTGCTTAATAACATTCGCAATCCAGCGCGGAGGTGGTGGCACCAACGTCAGCACCACTGGATTA 3291  
1236 GGAGTCGCTACGCGGCGCAACCCAGGAGTCTGCGCGCTAGTACCGCGCGATCGCAC 1295  
3292 GGAGTCGCTACGCGGCGCAACCCAGGAGTCTGCGCGCTAGTACCGCGCGATCGCAC 3351  
1296 GCATTGAAGAGTGGACGCACTGAAGAGAAATCGCGAAGAGCAGCGCGCGCAGCAGGCC 1355  
3352 GCATTGAAGAGTGGACGCACTGAAGAGAAATCGCGAAGAGCAGCGCGCGCAGCAGGCC 3411  
1356 GAGATGAAGAGGAGAGAGTGGCTGATGAACAGGATCGCGCAATCCAACTGGGAG 1415  
3412 GAGATGAAGAGGAGAGAGTGGCTGATGAACAGGATCGCGCAATCCAACTGGGAG 3471  
1416 ACGGCGCAATGATACGCGCAATATACAGACGCTGGAAATTTGTCCGCTGCTCGATGGC 1475

Db 3472 ACGGCGCAATGATACGCAATATACAGACGCTGGAATTTGTCCGCTGCTCGATGGC 3531  
QY 1476 CAGGCGTGCATGACCATCAGATCAGATCAGTGTGCGTGGCAGCGTCCATTTAGCGGCAAG 1535  
Db 3532 CAGGCGTGCATGACCATCAGATCAGATCAGTGTGCGTGGCAGCGTCCATTTAGCGGCAAG 3591  
QY 1536 GAGTCAATCGCAGGAGATCGATGTCATTTTCGGTGGCGGCAAGGACATCTCATCGTG 1595  
Db 3592 GAGTCAATCGCAGGAGATCGATGTCATTTTCGGTGGCGGCAAGGACATCTCATCGTG 3651  
QY 1596 CACGAGCGCGCAGCAAGGTCGACCTCACCAGTTCTCGAGAGAACCAAGTTTCGGTTC 1655  
Db 3652 CACGAGCGCGCAGCAAGGTCGACCTCACCAGTTCTCGAGAGAACCAAGTTTCGGTTC 3711  
QY 1656 GACTACGCTTCAACGACGACGTCGACATGCCATGGTATACAAATACACAGCAAGCC 1714  
Db 3712 GACTACGCTTCAACGACGACGTCGACATGCCATGGTATACAAATACCAATCC 3770

RESULT 6  
AC104510/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

AC104510 168583 bp DNA linear INV 13-JUN-2002  
Drosophila melanogaster X BAC RP98-26J12 (Roswell Park Cancer  
Institute Drosophila BAC Library) complete sequence.  
AC104510  
AC104510.6 GI:21328459  
HTG.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 168583)  
Muzny, D., Scherer, S., Adams, M.D., Holt, R.A., Evans, C.A.,  
Gocayne, J.D., Tabor, P., Williamson, A., Homsai, P.H.,  
Dugan-Rocha, S.D., Sodergren, E.S., Hodgson, A.H., Chen, R.C.,  
Ayele, M., Scott, G.S., Worley, K.W., Amaratides, P.G., Brandon, R.C.,  
Rogers, Y., An, H., Baldwin, D., Beeson, K.Y., Brown, M., Buhay, C.,  
Busam, D.A., Center, A., Chen, G., Chen, Z., Clerc-Blankenbush, K.,  
Davenport, L.B., Dietz, S.M., Ding, Y., Dodson, K., Doup, L.E.,  
Draper, H., Emery-Cohen, A., Ferreira, S., Garg, N.D.S., Houck, J.,  
Hostin, D., Howland, T.J., Hume, J., Ibegwam, C., Jalali, M., Kovar, C.,  
Liu, W., Mattei, B., McIntosh, T.C., Morgan, M., Moy, M., Murphy, B.,  
Nelson, K.A., Nasse, Y., Nguyen, N., Perez, L., Pittman, G.S., Puri, V.,  
Schaefer, F., Shen, H., Strong, R., Tector, C., Wang, Q., Williams, S.M.,  
Xiang, J., Zaveri, J.S., Zhou, J., Zorrilla, S., Smith, H.O.,  
Wheeler, D., Weinstock, G., Gibbs, R. and Venter, J.C.  
Direct Submission  
Unpublished  
2 (bases 1 to 168583)  
Worley, K.C., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amarutunge, H.C., Are, J.R., Banks, T., Barbara, J.,  
Benton, J., Bimige, K., Blankenburg, K., Bonnin, D., Bouck, J.,  
Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,  
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,  
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,  
Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,  
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,  
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,  
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,  
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,  
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,  
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,  
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,  
Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,  
Hollins, B., Homsai, P., Howard, S., Huber, J., Hulyk, S., Hume, J.,  
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, R., Korvah, J.,  
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,  
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, C., Liu, W.,  
Loulseged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,





Db	69229	TACAGGTGAGTCAAGAGGAAGGATCATGGACATGATTACGGTGGGGCAGAGCGTCAAG	69170
Qy	696	ATCAAGCGGACGAGTGGCGCGCTCCACATGGCGCGTGGTGGCGGTGATCAACCGTGGGC	755
Db	69169	ATCAAGCGGACGAGTGGCGCGCTCCACATGGCGCGTGGTGGCGGTGATCAACCGTGGGC	69110
Qy	756	AAAGTCATCAAGTGAATGTTACGAGCGCGCGGAAACGAAGCGGCAAGAGGAGGTAGAACTG	815
Db	69109	AAAGTCATCAAGTGAATGTTACGAGCGCGGAAACGAAGCGGCAAGAGGAGGTAGAACTG	69050
Qy	816	GACGCTACTCAGTCAATCCGAGTGAATGCAAGTACTGTGCAACAGCAGCCCGCC	875
Db	69049	GACGCTACTCAGTCAATCCGAGTGAATGCAAGTACTGTGCAACAGCAGCCCGCC	68990
Qy	876	CCGAGCGCAAGAACAGCAGCCCGCGGATGACCTTCGCTAAATCCACACAACTG	935
Db	68989	CCGAGCGCAAGAACAGCAGCCCGCGGATGACCTTCGCTAAATCCACACAACTG	68930
Qy	936	GCTATCGGTGGCAATCTCACACAGCGGTATGACCATGGCGGAAACATGCTGAACAAATC	995
Db	68929	GCTATCGGTGGCAATCTCACACAGCGGTATGACCATGGCGGAAACATGCTGAACAAATC	68870
Qy	996	CAGGAAAGCGAGTCCCAATCCGATTTGTGACGAGCAATAGCGTGAATCAAAACAGC	1055
Db	68869	CAGGAAAGCGAGTCCCAATCCGATTTGTGACGAGCAATAGCGTGAATCAAAACAGC	68810
Qy	1056	AACTCCACACTACCGCGCGGAGGTGGTGACACACAACTGACCAACCACTGGATTA	1115
Db	68809	AACTCCACACTACCGCGCGGAGGTGGTGACACACAACTGACCAACCACTGGATTA	68750
Qy	1116	CAGCGTCCACGCTACTCGCAAGCTGCTACCGGCGCAGCAGCAGCAAGATCGCTCGCG	1175
Db	68749	CAGCGTCCACGCTACTCGCAAGCTGCTACCGGCGCAGCAGCAGCAAGATCGCTCGCG	68690
Qy	1176	GTGCTTAATACATTCGCAATCCAGCGCGGAGCGAGTGTGTCGCGCGCACAA	1235
Db	68689	GTGCTTAATACATTCGCAATCCAGCGCGGAGCGAGTGTGTCGCGCGCACAA	68630
Qy	1236	GGAGTCGCACTCGCGGCGCAGCAGCGAGTGTGCGCGGTGTGACCGGCGATCGAC	1295
Db	68629	GGAGTCGCACTCGCGGCGCAGCAGCGAGTGTGCGCGGTGTGACCGGCGATCGAC	68570
Qy	1296	GCATTGAAGAGGTGGAGCGACTGAAGGAGATCCGAGAGCGAGCGCGCGCAGCGCC	1355
Db	68569	GCATTGAAGAGGTGGAGCGACTGAAGGAGATCCGAGAGCGAGCGCGCGCAGCGCC	68510
Qy	1356	GAGTGAAGAGGAGAGGTGGCGCTGATGAACAGAGATCCGCGCAATCCAAATCGGAG	1415
Db	68509	GAGTGAAGAGGAGAGGTGGCGCTGATGAACAGAGATCCGCGCAATCCAAATCGGAG	68450
Qy	1416	ACGGCGCAATCATACGCAATATCAGAGCAGCTGGAAATTTGTGCGCGTCTCGATGGC	1475
Db	68449	ACGGCGCAATCATACGCAATATCAGAGCAGCTGGAAATTTGTGCGCGTCTCGATGGC	68390
Qy	1476	CAGCGCGTCCGATGACCATCAGATCAGTGTGCGTGGCGCAAGCGTCCCATTAGCGCAAG	1535
Db	68389	CAGCGCGTCCGATGACCATCAGATCAGTGTGCGTGGCGCAAGCGTCCCATTAGCGCAAG	68330
Qy	1536	GAGTCAATCGAAGGAGATCGATGTCATTTGCGTGGCGCGCAAGGACATGCTCATCGTG	1595
Db	68329	GAGTCAATCGAAGGAGATCGATGTCATTTGCGTGGCGCGCAAGGACATGCTCATCGTG	68270
Qy	1596	CACGAGCGCGCAGAGGTGACCTCAGTTCAGTTCCTGGAGAACCAAGTTTGGCTTC	1655
Db	68269	CACGAGCGCGCAGAGGTGACCTCAGTTCAGTTCCTGGAGAACCAAGTTTGGCTTC	68210
Qy	1656	GACTACGCTTCAACGACAGCAGTGGCAATGCGATGCGTATCAATACACAGGCAAGCC	1714
Db	68209	GACTACGCTTCAACGACAGCAGTGGCAATGCGATGCGTATCAATACACAGGCAAGCC	68151

RESULT

AE003485/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AE003485 283075 bp DNA linear INV 10-AUG-2005  
Drosophila melanogaster chromosome X, section 37 of the complete sequence.  
AE003485 AE002593 AE014298  
AE003485.3 GI:22832075  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
AUTHORS

1 (bases 1 to 283075)  
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H., Blaise J., Blaise J., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L., Abril J.F., Agbayani A., An H.J., Andrews P., Binkov D., Baldwin D., Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter A., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W., Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalaali M., Kalush F., Karpen G.H., Ke Z., Kesterson J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Leake P., Lei Y., Levine A.A., Li J., Li J., Liang Y., Lin X., Liu X., Matheis B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier S., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Wasearman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage J., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Ye R.F., Zaveri J.S., Zhang M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M. and Venter J.C.  
The genome sequence of Drosophila melanogaster  
Science 287 (5461), 2185-2195 (2000)

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

2 (bases 1 to 283075)  
Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R., Paclet J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A. and Rubin G.M.  
Finishing a whole-genome shotgun: release 3 of the Drosophila melanogaster euchromatic genome sequence  
Genome Biol. 3 (12), RESEARCH0079 (2002)  
12537568

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

3 (bases 1 to 283075)  
Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M. and Lewis S.E.





Db	49722	TTGGGTTGCAATTTAGTTGTAATGCTTGAATTCGGTTCGGGGCTTAGTTTCCACAAGTT	49663	Db	48642	ACGGCGCAATGATACGCAATATACAGACGACGCTGGAATTTGTGCGCGCTCTCGATGGC	48583
Qy	396	TATCGCTCGTCAAGAAACAACAAATAAATTTATTTTCGACCTAAAAATCTGACTAAAT	455	Qy	1476	CAGCGCTCGATGACCATCAGATCACAGTGTGCGTGGCGCAAGCGTCCCATTTAGCGCGAAG	1535
Db	49662	TATCGCTCGTCAAGAAACAACAAATAAATTTATTTTCGACCTAAAAATCTGACTAAAT	49603	Db	48582	CAGCGCTCGATGACCATCAGATCACAGTGTGCGTGGCGCAAGCGTCCCATTTAGCGCGAAG	48523
Qy	456	TGTGTTTTTTTGTATTATTTATTTAGGCACATTTTGACACACCAACGCTAGTTACTAC	515	Qy	1536	GAGTCAATCCAAAGGAGATCGATGTCTATTTTCGGTGGCGGCAAGGACATGCTCATCGTG	1595
Db	49602	TGTGTTTTTTTGTATTATTTATTTAGGCACATTTTGACACACCAACGCTAGTTACTAC	49543	Db	48522	GAGTCAATCCAAAGGAGATCGATGTCTATTTTCGGTGGCGGCAAGGACATGCTCATCGTG	48463
Qy	516	ATCTACGACTAACGGAATCTCTCTGCAAGCAGTGGAAATGTTGCTGTCCATCAAGCAGTAC	575	Qy	1596	CACGAGCGCGCACCAAGGTGACCTCACCAAGTTCTCTGGAGAACCAAGTTTCGCTTC	1655
Db	49542	ATCTACGACTAACGGAATCTCTCTGCAAGCAGTGGAAATGTTGCTGTCCATCAAGCAGTAC	49483	Db	48462	CACGAGCGCGCACCAAGGTGACCTCACCAAGTTCTCTGGAGAACCAAGTTTCGCTTC	48403
Qy	576	TCGGAGTTAACCCAGGATTAAGCCGGAGNAAGAGAGATCGGTGGAGATAGAGATA	635	Qy	1656	GACTACGCTTCAACGACGATGCGACATGCCATGTTATCAAAATACACAGCCAAAGCC	1714
Db	49482	TCGGAGTTTAACCCAGGATTAAGCCGGAGNAAGAGAGATCGGTGGAGATAGAGATA	49423	Db	48402	GACTACGCTTCAACGACGATGCGACATGCCATGTTATCAAAATGAGTATCAATCC	48344
Qy	636	TACAGTGGAGTCAAGAGAAAGGATCATGGATGATTACGGTGGCGGAGAGCGTCAAG	695	RESULT 8			
Db	49422	TACAGTGGAGTCAAGAGAAAGGATCATGGATGATTACGGTGGCGGAGAGCGTCAAG	49363	AC116537			
Qy	696	ATCAAGCGGAACGATGGCGGCTCCACATGCGCGTGGTGGCGGTGATCAACCAAGTCGGGC	755	LOCUS			
Db	49362	ATCAAGCGGAACGATGGCGGCTCCACATGCGCGTGGTGGCGGTGATCAACCAAGTCGGGC	49303	DEFINITION	AC116537	334028 bp DNA linear HTG 18-MAY-2002	
Qy	756	AGTGCATCACAGTCCGATGTTACAGCGCGCGGAAACGAAGGCAAGAGGTAGAACTG	815	ACCESSION	AC116537.2	GI:20898595	
Db	49302	AGTGCATCACAGTCCGATGTTACAGCGCGCGGAAACGAAGGCAAGAGGTAGAACTG	49243	VERSION	AC116537	HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.	
Qy	816	GACGCCATACCTCAAGTCCGAGGCTAATGCAAGATATCTGTCGAACAGCAGCCGCGCC	875	KEYWORDS	AC116537	HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.	
Db	49242	GACGCCATACCTCAAGTCCGAGGCTAATGCAAGATATCTGTCGAACAGCAGCCGCGCC	49183	SOURCE	AC116537	Drosophila melanogaster (fruit fly)	
Qy	876	CGGAGCCCAAGAAACAAGCCACCGCGCGATGAACCTCTCGCTTAATCCCAACAATCG	935	ORGANISM	AC116537	Drosophila melanogaster	
Db	49182	CGGAGCCCAAGAAACAAGCCACCGCGCGATGAACCTCTCGCTTAATCCCAACAATCG	49123	REFERENCE		1 (bases 1 to 334028)	
Qy	936	GCTATCGGTGCAATCTCACCAGCGTATGACCATGGCGGGAACATCTGAACAGATC	995	AUTHORS		Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,	
Db	49122	GCTATCGGTGCAATCTCACCAGCGTATGACCATGGCGGGAACATCTGAACAGATC	49063			Alsbrooks, S.L., Amarutunge, H.C., Are, J.R., Ayele, M., Banks, T.,	
Qy	996	CAGGAAAGCCAGTCCGATTTCCCAATCCGATTTGTACAGCAATAGGTTGAAATACAAACAGC	1055			Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,	
Db	49062	CAGGAAAGCCAGTCCGATTTCCCAATCCGATTTGTACAGCAATAGGTTGAAATACAAACAGC	49003			Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,	
Qy	1056	AACTCCAACTACGCGCGGAGGTGGTGGCCACCAACGTCGACCACTGGATTA	1115			Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,	
Db	49002	AACTCCAACTACGCGCGGAGGTGGTGGCCACCAACGTCGACCACTGGATTA	48943			Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,	
Qy	1116	CAGCGTCCACGCTACTCCGACGCTGCTACCGGCGAGCAGACAGGATCGCTCGCG	1175			Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,	
Db	48942	CAGCGTCCACGCTACTCCGACGCTGCTACCGGCGAGCAGACAGGATCGCTCGCG	48883			Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,	
Qy	1176	GTGCTTAATACACATTCGCCAATCCAGCGCGGAGCGAGTGTGTCGGCGGCAAA	1235			Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,	
Db	48882	GTGCTTAATACACATTCGCCAATCCAGCGCGGAGCGAGTGTGTCGGCGGCAAA	48823			Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,	
Qy	1236	GGAGTGGCACTGCGGCGCACCAACGAGGAGTGGCGCGTGTAGTACCCGCGCATCGCAC	1295			Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,	
Db	48822	GGAGTGGCACTGCGGCGCACCAACGAGGAGTGGCGCGTGTAGTACCCGCGCATCGCAC	48763			Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,	
Qy	1296	GCATTTAAGAGGTGGAGGACTGAGGAGATCGCGAGAGCGACGCGCGGACAGGCC	1355			Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,	
Db	48762	GCATTTAAGAGGTGGAGGACTGAGGAGATCGCGAGAGCGACGCGCGGACAGGCC	48703			Garrett, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,	
Qy	1356	GAGATGAAGGAGGAGAGGTGGCGCTGATGAACAGAGTCCGGGCAATCCAACTGGAG	1415			Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,	
Db	48702	GAGATGAAGGAGGAGAGGTGGCGCTGATGAACAGAGTCCGGGCAATCCAACTGGAG	48643			Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,	
Qy	1416	ACGGCGCAATGATACCGCAATATACAGACGACGTTGGAATTTGTGCGCGCTCTCGATGGC	1475			Honsi, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,	
						Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,	
						Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,	
						Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,	
						Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Loulseg, H.,	
						Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,	
						Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,	
						Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,	
						Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,	
						Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,	
						Nguyen, N., Nickerson, E., Nwokkwo, S., Ogih, M., Okwionu, G.,	
						Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,	
						Peters, L., Pichens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,	
						Rives, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Savery, G.,	
						Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,	
						Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H.,	
						Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,	
						Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,	
						Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,	
						Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,	
						Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K.,	
						Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,	
						Weinstock, G. and Gibbs, R.	

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Direct Submission  
Unpublished  
2 (bases 1 to 334028)  
Worley, K.C.  
Direct Submission  
Submitted (29-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 334028)  
Worley, K.C.  
Direct Submission  
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 17, 2002 this sequence version replaced gi:19807669.

## COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: DRIT  
Center clone name: RP98-9F14  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 327497 bases at least Q40  
Consensus quality: 328682 bases at least Q30  
Consensus quality: 329600 bases at least Q20  
Estimated insert size: 175224; sum-of-contrigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contrigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 14 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2752: contig of 2752 bp in length  
\* 2753 2852: gap of unknown length  
\* 2853 9484: contig of 6632 bp in length  
\* 9484 9584: gap of unknown length  
\* 9585 13752: contig of 4168 bp in length  
\* 13753 13852: gap of unknown length  
\* 13853 30299: contig of 16447 bp in length  
\* 30300 30399: gap of unknown length  
\* 30400 45744: contig of 15345 bp in length  
\* 45745 45844: gap of unknown length  
\* 45845 59632: contig of 13788 bp in length  
\* 59633 59732: gap of unknown length  
\* 59733 80931: contig of 21199 bp in length  
\* 80932 81031: gap of unknown length  
\* 81032 104548: contig of 23517 bp in length  
\* 104549 104648: gap of unknown length  
\* 104649 123841: contig of 19193 bp in length  
\* 123842 123941: gap of unknown length  
\* 123942 145561: contig of 21620 bp in length  
\* 145562 145661: gap of unknown length  
\* 145662 183677: contig of 38016 bp in length  
\* 183678 183777: gap of unknown length  
\* 183778 21275: contig of 33498 bp in length  
\* 21276 21375: gap of unknown length  
\* 21376 268427: contig of 51052 bp in length  
\* 268428 268527: gap of unknown length  
\* 268528 334028: contig of 65501 bp in length.

FEATURES  
Location/Qualifiers  
1. .334028

/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/chromosome="X"  
/clone="RP98-9F14"

2753..2852  
/estimated\_length=unknown  
9485..9584  
/estimated\_length=unknown  
13753..13852  
/estimated\_length=unknown  
30300..30399  
/estimated\_length=unknown  
45745..45844  
/estimated\_length=unknown  
59633..59732  
/estimated\_length=unknown  
80932..81031  
/estimated\_length=unknown  
104549..104648  
/estimated\_length=unknown  
123842..123941  
/estimated\_length=unknown  
145562..145661  
/estimated\_length=unknown  
183678..183777  
/estimated\_length=unknown  
21276..21375  
/estimated\_length=unknown  
268428..268527  
/estimated\_length=unknown

## ORIGIN

Query Match 50.2%; Score 1547.8; DB 14; Length 334028;  
Best Local Similarity 99.6%; Pred. No. 2.8e-286; Indels 0; Gaps 0;  
Matches 1552; Conservative 0; Mismatches 7;  
QY 156 AGCCAGCGGCTGCGGTTTCTCTTTTGTATCGGTTGCCAGAGCGCAACGAAACGACA 215  
DB 217943 AGCCAGCGGCTGCGGTTTCTCTTTTGTATCGGTTGCCAGAGCGCAACGAAACGACA 218002  
QY 216 GTTGGCATTTGAATTCAGCACAACACACACATCTAACGCCGCCGCCGAGCAGCACAC 275  
DB 218003 GTTGGCATTTGAATTCAGCACAACACACATCTAACGCCGCCGCCGAGCAGCACAC 218062  
QY 276 ACACACACTGGGACACTCGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 335  
DB 218063 ACACACACTGGGACACTCGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 218122  
QY 336 TTGGGTTTCGATTTAGTTTGTCAATGCTTGAATTCGGTTTCGGGGCTTAGTTTCCACAAGTT 395  
DB 218123 TTGGGTTTCGATTTAGTTTGTCAATGCTTGAATTCGGTTTCGGGGCTTAGTTTCCACAAGTT 218182  
QY 396 TATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCGACTTAAATAATCTGACTTAAAT 455  
DB 218183 TATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCGACTTAAATAATCTGACTTAAAT 218242  
QY 456 TGTGTTTTTTTGTATGTTATTTATTTAGGCACATTTTCGACACCAACCAAGCTAGTTACTAC 515  
DB 218243 TGTGTTTTTTTGTATGTTATTTATTTAGGCACATTTTCGACACCAACCAAGCTAGTTACTAC 218302  
QY 516 ATCTACGACTAACCGAAATCTCTCTCGAAGCAGTGGAAAGTTCTGTCTCAATCAAGCAGTAC 575  
DB 218303 ATCTACGACTAACCGAAATCTCTCTCGAAGCAGTGGAAAGTTCTGTCTCAATCAAGCAGTAC 218362  
QY 576 TCGGAGTTTAAACGACGATTAAGCCGGGAGAAAGAGAAAGAGATCGGTGGGAGAAATAGATATA 635  
DB 218363 TCGGAGTTTAAACGACGATTAAGCCGGGAGAAAGAGAAAGAGATCGGTGGGAGAAATAGATATA 218422  
QY 636 TACAGTGGAGTCAAGAGGAGGATCATGCATGATTACGTTGGGCGGAGCGCTCAAG 695  
DB 218423 TACAGTGGAGTCAAGAGGAGGATCATGCATGATTACGTTGGGCGGAGCGCTCAAG 218482

```
QY 696 ATCAAGCGGACGAGTGGCGCGCTCCACATGGCGCTGGTGGCGGTGATCAACCAAGTCGGGC 755
Db 218483 ATCAAGCGGACGAGTGGCGCGCTCCACATGGCGCTGGTGGCGGTGATCAACCAAGTCGGGC 218542
QY 756 AAGTGCATCAGTGCAGTGGTACGAGCGCGCGGAAACGAAGGGCAAGAGGTAGAACTG 815
Db 218543 AAGTGCATCAGTGCAGTGGTACGAGCGCGCGGAAACGAAGGGCAAGAGGTAGAACTG 218602
QY 816 GAGCGCATCTACGCTCAATCCGAGCTAATGCAAGTAAATGCAAGTACTGTGCAACGACGCGGCC 875
Db 218603 GAGCGCATCTACGCTCAATCCGAGCTAATGCAAGTAAATGCAAGTACTGTGCAACGACGCGGCC 218662
QY 876 CCGGAGCCCAAGAAACAGCCACCGCGCGATGAACCTCTCCGCTAAATCCCAACACAATCG 935
Db 218663 CCGGAGCCCAAGAAACAGCCACCGCGCGATGAACCTCTCCGCTAAATCCCAACACAATCG 218722
QY 936 GTATCGTGGCAATCTACAGCCGCTATGACCATGGCGGAAACATGCTGAACCAAGATC 995
Db 218723 GTATCGTGGCAATCTACAGCCGCTATGACCATGGCGGAAACATGCTGAACCAAGATC 218782
QY 996 CAGGAAAGCCAGTGCATTCCTCAATCCGATTTGTCAGCAGCAATAGCGTGAATACAAACAGC 1055
Db 218783 CAGGAAAGCCAGTGCATTCCTCAATCCGATTTGTCAGCAGCAATAGCGTGAATACAAACAGC 218842
QY 1056 AACTCCAACTACGCGCGCGGAGGTGGTGCCACCAACGTCGACGACCACTGGATTA 1115
Db 218843 AACTCCAACTACGCGCGCGGAGGTGGTGCCACCAACGTCGACGACCACTGGATTA 218902
QY 1116 CAGCGTCAACGCTACTCGGAGCTCTACCGGCCAGCAGCAGCAAGGATCGCTCGGCG 1175
Db 218903 CAGCGTCAACGCTACTCGGAGCTCTACCGGCCAGCAGCAGCAAGGATCGCTCGGCG 218962
QY 1176 GTGCGCTAATAACACATTCGCCAATCCAGCGCGGAGCGAGTGTGCTGGTCCGGCGGCACAA 1235
Db 218963 GTGCGCTAATAACACATTCGCCAATCCAGCGCGGAGCGAGTGTGCTGGTCCGGCGGCACAA 219022
QY 1236 GAGTGCACCTCGCGGCCAACACCCAGGAGTGTGCGCGCTAGTACCCGCGATCGCAC 1295
Db 219023 GAGTGCACCTCGCGGCCAACACCCAGGAGTGTGCGCGCTAGTACCCGCGATCGCAC 219082
QY 1296 GCATTGAAAGAGTGGAGCACTGAAGGAGAAATCGCGAAGACGACGCGCCGACAGGCC 1355
Db 219083 GCATTGAAAGAGTGGAGCACTGAAGGAGAAATCGCGAAGACGACGCGCCGACAGGCC 219142
QY 1356 GAGATGAAGGAGGAGAGTGCCTGATGACAGGATCCGGCAATCCAAATCGGAG 1415
Db 219143 GAGATGAAGGAGGAGAGTGCCTGATGACAGGATCCGGCAATCCAAATCGGAG 219202
QY 1416 ACGGCGCAATGATACGCGCAATATCAGAGCAGCTGGAAATTTGTGCCCTGCTCGATGGC 1475
Db 219203 ACGGCGCAATGATACGCGCAATATCAGAGCAGCTGGAAATTTGTGCCCTGCTCGATGGC 219262
QY 1476 CAGGCGCTGATGACCATCAGATCAGATGCTGGTGGCAAGCGTCCCAATAGCGCAAG 1535
Db 219263 CAGGCGCTGATGACCATCAGATCAGATGCTGGTGGCAAGCGTCCCAATAGCGCAAG 219322
QY 1536 GAGGTCAATCCAGGAGATCGATGCTATTTGCTGGCGCGCAAGGACATGCTCATCGTG 1595
Db 219323 GAGGTCAATCCAGGAGATCGATGCTATTTGCTGGCGCGCAAGGACATGCTCATCGTG 219382
QY 1596 CACGAGCGCGCAGAGGTGCACCTCACCAAGTTCTGGAGAACCAAGTTTCGTTTC 1655
Db 219383 CACGAGCGCGCAGAGGTGCACCTCACCAAGTTCTGGAGAACCAAGTTTCGTTTC 219442
QY 1656 GACTACGCTTCAAGCAGCAGTGGCAATGCCATGGTATCAATCAACACAGCAAGCC 1714
Db 219443 GACTACGCTTCAAGCAGCAGTGGCAATGCCATGGTATCAATCAACACAGCAAGCC 219501
RESULT 8
AC007521
LOCUS 180919 bp DNA linear HTG 15-MAR-2002
DEFINITION Drosophila melanogaster strain y; cn bw sp chromosome X clone
```

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BACR49A04 (D698) RPCI-98 49 A.4 map 10A2-10B2, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 61 unordered pieces.  
AC007521 GI:19483792  
HTG, HTGS PHASE1.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
AUTHORS

1 (bases 1 to 180919)  
Celisner, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,  
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciegiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,  
Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,  
Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B.,  
Richards, S., Sethi, H., Svirskas, R.R., Wan, K.H., Webster, D.,  
Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.  
Sequencing of *Drosophila melanogaster*  
Unpublished

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

2 (bases 1 to 180919)  
Celisner, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,  
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciegiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,  
Kim, E., Lee, B., Lewis, S., Li, P., Lotman, M.A., Mazda, P.,  
Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,  
Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,  
Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and  
Rubin, G.M.  
Direct Submission  
Submitted (10-MAY-1999) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA  
3 (bases 1 to 180919)  
Celisner, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,  
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciegiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,  
Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,  
Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B.,  
Richards, S., Sethi, H., Svirskas, R.R., Wan, K.H., Webster, D.,  
Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.  
Direct Submission  
Submitted (15-MAR-2002) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

On Mar 15, 2002 this sequence version replaced gi:7191020.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive Web site (<http://www.fruitfly.org/sequence/>) or send email  
to [bdgp@fruitfly.berkeley.edu](mailto:bdgp@fruitfly.berkeley.edu). All contigs in this submission meet  
the following cutoffs: length >= 200 bases.  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 61 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

TITLE  
JOURNAL  
COMMENT

717: contig of 717 bp in length  
797: gap of unknown length  
1345: contig of 548 bp in length  
1425: gap of unknown length  
2041: contig of 616 bp in length  
2042: gap of unknown length  
2122: contig of 564 bp in length  
2685: gap of unknown length  
2765: contig of 743 bp in length  
3588: gap of unknown length  
4723: contig of 1135 bp in length  
4803: gap of unknown length  
5781: contig of 978 bp in length

```
* 5782 5861: gap of unknown length
* 5962 7147: contig of 1286 bp in length
* 7148 7227: gap of unknown length
* 7228 9276: contig of 2049 bp in length
* 9357 9356: gap of unknown length
* 10427 10426: contig of 1070 bp in length
* 10507 11545: contig of 1039 bp in length
* 11546 11635: gap of unknown length
* 11636 12578: contig of 953 bp in length
* 12579 12658: gap of unknown length
* 12659 13814: contig of 1156 bp in length
* 13815 13894: gap of unknown length
* 13895 15132: contig of 1238 bp in length
* 15133 15212: gap of unknown length
* 15213 16663: contig of 1451 bp in length
* 16664 16743: gap of unknown length
* 16744 17742: contig of 999 bp in length
* 17743 17822: gap of unknown length
* 17823 19718: contig of 1896 bp in length
* 19719 19798: gap of unknown length
* 19799 21740: contig of 1942 bp in length
* 21741 21820: gap of unknown length
* 21821 24129: contig of 2309 bp in length
* 24130 24209: gap of unknown length
* 24210 26333: contig of 2124 bp in length
* 26334 26413: gap of unknown length
* 26414 27843: contig of 1430 bp in length
* 27844 27923: gap of unknown length
* 27924 29790: contig of 1867 bp in length
* 29791 29870: gap of unknown length
* 29871 33117: contig of 3247 bp in length
* 33118 33197: gap of unknown length
* 33198 36063: contig of 2866 bp in length
* 36064 36143: gap of unknown length
* 36144 39335: contig of 3192 bp in length
* 39336 39415: gap of unknown length
* 39416 43414: contig of 3999 bp in length
* 43415 43494: gap of unknown length
* 43495 46045: contig of 2551 bp in length
* 46046 46125: gap of unknown length
* 46126 50185: contig of 4060 bp in length
* 50186 50265: gap of unknown length
* 50266 55215: contig of 4950 bp in length
* 55216 55295: gap of unknown length
* 55296 59535: contig of 4240 bp in length
* 59536 59615: gap of unknown length
* 59616 65398: contig of 5783 bp in length
* 65399 65478: gap of unknown length
* 65479 73776: contig of 8298 bp in length
* 73777 73856: gap of unknown length
* 73857 78055: contig of 4199 bp in length
* 78056 78135: gap of unknown length
* 78136 90995: contig of 12860 bp in length
* 90996 91075: gap of unknown length
* 91076 102049: contig of 10974 bp in length
* 102050 102129: gap of unknown length
* 102130 111675: contig of 9546 bp in length
* 111676 111755: gap of unknown length
* 111756 122893: contig of 11138 bp in length
* 122894 122973: gap of unknown length
* 122974 139619: contig of 16646 bp in length
* 139620 139699: gap of unknown length
* 139700 166353: contig of 26654 bp in length
* 166354 166433: gap of unknown length
* 166434 166864: contig of 431 bp in length
* 166865 166944: gap of unknown length
* 166945 167735: contig of 791 bp in length
* 167736 167815: gap of unknown length
* 167816 168509: contig of 694 bp in length
* 168510 168589: gap of unknown length
* 168590 168928: contig of 339 bp in length
* 168929 169008: gap of unknown length

* 169009 169748: contig of 740 bp in length
* 169749 169828: gap of unknown length
* 169829 170148: contig of 320 bp in length
* 170149 170228: gap of unknown length
* 170229 170610: contig of 382 bp in length
* 170611 170890: gap of unknown length
* 170891 170927: contig of 237 bp in length
* 170928 171007: gap of unknown length
* 171008 171769: contig of 762 bp in length
* 171770 171849: gap of unknown length
* 171850 172190: contig of 341 bp in length
* 172191 172270: gap of unknown length
* 172271 172592: contig of 322 bp in length
* 172593 172672: gap of unknown length
* 172673 173139: contig of 467 bp in length
* 173140 173219: gap of unknown length
* 173220 175301: contig of 2082 bp in length
* 175302 175381: gap of unknown length
* 175382 175850: contig of 469 bp in length
* 175851 175930: gap of unknown length
* 175931 176376: contig of 446 bp in length
* 176377 176456: gap of unknown length
* 176457 176939: contig of 483 bp in length
* 176940 177019: gap of unknown length
* 177020 177549: contig of 530 bp in length
* 177550 177629: gap of unknown length
* 177630 178146: contig of 517 bp in length
* 178147 178226: gap of unknown length
* 178227 178600: contig of 634 bp in length
* 178601 178940: gap of unknown length
* 178941 179549: contig of 609 bp in length
* 179550 179629: gap of unknown length
* 179630 180203: contig of 574 bp in length
* 180204 180283: gap of unknown length
* 180284 180319: contig of 636 bp in length.

FEATURES
      source
      1..180919
      /organism="Drosophila melanogaster"
      /mol_type="genomic DNA"
      /strain="Y; cn bw sp"
      /db_xref="taxon:7227"
      /chromosome="X"
      /map="10A2-10B2"
      /clone="BACR49A04 (D698) RPCI-98 49.A.4"
      /clone_lib="RPCI-98 (Roswell Park Cancer Institute
      Drosophila melanogaster BAC library, partial EcoRI in
      pBACe3.6"
      718...797
      /estimated_length=unknown
      1346...1425
      /estimated_length=unknown
      2042...2121

Query Match      36.3%; Score 1117.6; DB 14; Length 180919;
Best Local Similarity 93.0%; Pred. NO. 9.2e-204;
Matches 1131; Conservative 0; Mismatches 84; Indels 1; Gaps 1;

QY 156 AGCCAGCGGGCTGCGGTTTTTCCCTTTTGTATCGTTGCCAGCGCAACGAAACGACA 215
      |||||||
Db 110544 AGCCAGCGGGCTGCGGTTTTTCCCTTTTGTATCGTTGCCAGCGCAACGAAACGACA 110603
      |||||||

QY 216 GTTGGCATTTCGAATTCAGCACAACACATACATAACCGCCGACCGCAAGCAGCACAC 275
      |||||||
Db 110604 GTTGGCATTTCGAATTCAGCACAACACATACATAACCGCCGACCGCAAGCAGCACAC 110663
      |||||||

QY 276 ACACACACTGGGACACTCGAAAAAAGAAAAAAGACGCTGTGGCGCACTCGACAAGCAG 335
      |||||||
Db 110664 ACACACACTGGGACACTCGAAAAAAGAAAAAAGACGCTGTGGCGCACTCGACAAGCAG 110723
      |||||||

QY 336 TTGGGTTTCGATTAGTTGTCATCGCTTGAATTCGGTTTCGGGGCTTAGTTCCCAAGTT 395
      |||||||
Db 110724 TTGGGTTTCGATTAGTTGTCATCGCTTGAATTCGGTTTCGGGGCTTAGTTCCCAAGTT 110783
      |||||||
```

```
QY 396 TATCGCTCGTCAAGAAACCAAGAAATAAATAATTTATTTTCGACCTAAAAATCTGACTAAT 455
Db 110784 TATCGCTCGTCAAGAAACCAAGAAATAAATAATTTATTTTCGACCTAAAAATCTGACTAAT 110843
QY 456 TGTGTTTTTTTGTATGTATTTATTTAGGCACATTTTGCACACCAACGCTAGTTACTTAC 515
Db 110844 TGTGTTTTTTTGTATGTATTTATTTAGGCACATTTTGCACACCAACGCTAGTTACTTAC 110903
QY 516 ATCTACGACTAACGGAATCTCTCTGCAAGCAGTGGAAATGTTCTCTCATCAAGCAGTAC 575
Db 110904 ATCTACGACTAACGGAATCTCTCTGCAAGCAGTGGAAATGTTCTCTCATCAAGCAGTAC 110963
QY 576 TCGGAGTTAAACGACGATAAGCCGGAGAAAGAGAAAGAGATCGGTGGAGAAATAGAGATA 635
Db 110964 TCGGAGTTAAACGACGATAAGCCGGAGAAAGAGAAAGAGATCGGTGGAGAAATAGAGATA 111023
QY 636 TACAGGTGGAGTCAAGAGGAAGGATCATGACATGATTAGGTGGGCGCAGAGGTCAAG 695
Db 111024 TACAGGTGGAGTCAAGAGGAAGGATCATGACATGATTAGGTGGGCGCAGAGGTCAAG 111083
QY 696 ATCAAGCGGACGATGGCGGCTCCACATGCGCGTGGTGGCGGTGATCAACCAAGTCGGGC 755
Db 111084 ATCAAGCGGACGATGGCGGCTCCACATGCGCGTGGTGGCGGTGATCAACCAAGTCGGGC 111143
QY 756 AAGTGCATCACAGTCGAATGGTACGAGCGCGCGCAAGCAAGGGCAAGAGGTAGAACTG 815
Db 111144 AAGTGCATCACAGTCGAATGGTACGAGCGCGCGCAAGCAAGGGCAAGAGGTAGAACTG 111203
QY 816 GAGGCATACCTACGCTCAATCCGAGGTATGCAAGATATCTGTGCAACAGCAGCGCGCC 875
Db 111204 GAGGCATACCTACGCTCAATCCGAGGTATGCAAGATATCTGTGCAACAGCAGCGCGCC 111263
QY 876 CGGAGGCCAAGAAACCAAGCAACCGCGCGATGACCTCTCGCTTAATCCACACATCG 935
Db 111264 CGGAGGCCAAGAAACCAAGCAACCGCGCGATGACCTCTCGCTTAATCCACACATCG 111323
QY 936 GCTATCGGTGCAATCTCACCAGCGGTATGACCATGCGCGGAAACATGCTGAACAAAGATC 995
Db 111324 GCTATCGGTGCAATCTCACCAGCGGTATGACCATGCGCGGAAACATGCTGAACAAAGATC 111383
QY 996 CAGNAAGCCAGTGTATCCCAATCCGATTTGTGACGAGCAATAGGTGAATACAAACAGC 1055
Db 111384 CAGNAAGCCAGTGTATCCCAATCCGATTTGTGACGAGCAATAGGTGAATACAAACAGC 111443
QY 1056 AACTCCACACTACGCGCGGAGGTGTGCGACCAACGTCGACGACCACTGGATTA 1115
Db 111444 AACTCCACACTACGCGCGGAGGTGTGCGACCAACGTCGACGACCACTGGATTA 111503
QY 1116 CAGGTCACCGGTACTCGCAAGCTGCTACCGGCGCAGCAGACAAAGATCGCTCGGCG 1175
Db 111504 CAGGTCACCGGTACTCGCAAGCTGCTACCGGCGCAGCAGACAAAGATCGCTCGGCG 111563
QY 1176 GTGCTTAATACATTTGCCAATCCAGCGCGGAGCAGT-GCTGTTCGGCGGACACA 1234
Db 111564 GTGCTTAATACATTTGCCAATCCAGCGCGGAGCAGTGTGCTGTGTCGGCGGACACA 111623
QY 1235 AGGAGTCCCACTCGGCGCACACCCAGGAGCTGGCGGCTAGTACCCGCGCATCGCA 1294
Db 111624 AGGAGTCCCACTCGGCGCACACCCAGGAGCTGGCGGCTAGTACCCGCGCGNNNNN 111683
QY 1295 CGCATTTGAAAGAGGTGGAGCGACTGAAGGAATTCGCGAGAAGCGACGCGCCGACAGGC 1354
Db 111684 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 111743
QY 1355 CGAGATGAAGAGAG 1370
Db 111744 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 111759
RESULT 10
AC007521/c 180919 bp DNA linear HTG 15-MAR-2002
LOCUS Drosophila melanogaster strain y; cn bw sp chromosome X clone
DEFINITION
```

```
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1717: contig of 717 bp in length
797: gap of unknown length
718
798
1345: contig of 548 bp in length
1346
1425: gap of unknown length
1426
2041: contig of 616 bp in length
2042
2121: gap of unknown length
2122
2685: contig of 564 bp in length
2686
2765: gap of unknown length
2766
3508: contig of 743 bp in length
3509
3588: gap of unknown length
3589
4723: contig of 1135 bp in length
4724
4803: gap of unknown length
4804
5781: contig of 978 bp in length
```

BACK49A04 (D698) RPCI-98 49.A.4 map 10A2-10B2, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 61 unordered pieces.

AC007521 GI:19483792

HTG, HTGS PHASE1.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.

Sequencing of Drosophila melanogaster

Unpublished

2 (bases 1 to 180919)

Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.

Direct Submission

Submitted (10-MAY-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

3 (bases 1 to 180919)

Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.

Direct Submission

Submitted (15-MAR-2002) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Mar 15, 2002 this sequence version replaced gi:7191020.

For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to [bdg@fruitfly.berkeley.edu](mailto:bdg@fruitfly.berkeley.edu). All contigs in this submission meet the following cutoffs: length >= 200 bases.

\* NOTE: This is a 'working draft' sequence. It currently consists of 61 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1

717: contig of 717 bp in length

797: gap of unknown length

718

798

1345: contig of 548 bp in length

1346

1425: gap of unknown length

1426

2041: contig of 616 bp in length

2042

2121: gap of unknown length

2122

2685: contig of 564 bp in length

2686

2765: gap of unknown length

2766

3508: contig of 743 bp in length

3509

3588: gap of unknown length

3589

4723: contig of 1135 bp in length

4724

4803: gap of unknown length

4804

5781: contig of 978 bp in length



```
* 5782 5861: gap of unknown length
* 5862 7147: contig of 1286 bp in length
* 7148 7227: gap of unknown length
* 7228 9276: contig of 2049 bp in length
* 9277 9356: gap of unknown length
* 9357 10426: contig of 1070 bp in length
* 10427 10506: gap of unknown length
* 10507 11545: contig of 1039 bp in length
* 11546 11625: gap of unknown length
* 11626 12578: contig of 953 bp in length
* 12579 13814: contig of 1156 bp in length
* 13815 13894: gap of unknown length
* 13895 15132: contig of 1238 bp in length
* 15133 15212: gap of unknown length
* 15213 16663: contig of 1451 bp in length
* 16664 16743: gap of unknown length
* 16744 17742: contig of 999 bp in length
* 17743 17822: gap of unknown length
* 17823 19718: contig of 1896 bp in length
* 19719 19798: gap of unknown length
* 19799 21740: contig of 1942 bp in length
* 21741 21820: gap of unknown length
* 21821 24129: contig of 2309 bp in length
* 24130 24209: gap of unknown length
* 24210 26333: contig of 2124 bp in length
* 26334 26413: gap of unknown length
* 26414 27843: contig of 1430 bp in length
* 27844 27923: gap of unknown length
* 27924 29790: contig of 1867 bp in length
* 29791 29870: gap of unknown length
* 29871 33117: contig of 3247 bp in length
* 33118 33197: gap of unknown length
* 33199 36063: contig of 2866 bp in length
* 36064 36143: gap of unknown length
* 36144 39335: contig of 3192 bp in length
* 39336 39415: gap of unknown length
* 39416 43414: contig of 3999 bp in length
* 43415 43494: gap of unknown length
* 43495 46045: contig of 2551 bp in length
* 46046 46125: gap of unknown length
* 46126 50185: contig of 4060 bp in length
* 50186 50265: gap of unknown length
* 50266 55215: contig of 4950 bp in length
* 55216 55295: gap of unknown length
* 55296 59535: contig of 4240 bp in length
* 59536 59615: gap of unknown length
* 59616 65398: contig of 5783 bp in length
* 65399 65478: gap of unknown length
* 65479 73776: contig of 8298 bp in length
* 73777 73856: gap of unknown length
* 73857 78055: contig of 4199 bp in length
* 78056 78135: gap of unknown length
* 78136 90995: contig of 12860 bp in length
* 90996 91075: gap of unknown length
* 91076 102050: contig of 10974 bp in length
* 102051 102129: gap of unknown length
* 102130 111675: contig of 9546 bp in length
* 111676 111756: gap of unknown length
* 111757 122893: contig of 11138 bp in length
* 122894 122973: gap of unknown length
* 122974 139619: contig of 16646 bp in length
* 139620 139699: gap of unknown length
* 139700 166353: contig of 26654 bp in length
* 166354 166433: gap of unknown length
* 166434 166864: contig of 431 bp in length
* 166865 166945: gap of unknown length
* 166946 167735: contig of 791 bp in length
* 167736 167815: gap of unknown length
* 167816 168509: contig of 694 bp in length
* 168510 168589: gap of unknown length
* 168590 168928: contig of 339 bp in length
* 168929 169008: gap of unknown length

* 169009 169748: contig of 740 bp in length
* 169749 169828: gap of unknown length
* 169829 170148: contig of 320 bp in length
* 170149 170228: gap of unknown length
* 170229 170610: contig of 382 bp in length
* 170611 170890: gap of unknown length
* 170891 170927: contig of 237 bp in length
* 170928 171007: gap of unknown length
* 171008 171769: contig of 762 bp in length
* 171770 171849: gap of unknown length
* 171850 172190: contig of 341 bp in length
* 172191 172270: gap of unknown length
* 172271 172592: contig of 322 bp in length
* 172593 172672: gap of unknown length
* 172673 173139: contig of 467 bp in length
* 173140 173219: gap of unknown length
* 173220 175301: contig of 2082 bp in length
* 175302 175381: gap of unknown length
* 175382 175850: contig of 469 bp in length
* 175851 175930: gap of unknown length
* 175931 176376: contig of 446 bp in length
* 176377 176456: gap of unknown length
* 176457 176939: contig of 483 bp in length
* 176940 177019: gap of unknown length
* 177020 177549: contig of 530 bp in length
* 177550 177629: gap of unknown length
* 177630 178146: contig of 517 bp in length
* 178147 178226: gap of unknown length
* 178227 178860: contig of 634 bp in length
* 178861 178940: gap of unknown length
* 178941 179549: contig of 609 bp in length
* 179550 179629: gap of unknown length
* 179630 180203: contig of 574 bp in length
* 180204 180283: gap of unknown length
* 180284 180919: contig of 636 bp in length.

FEATURES
      source
      1..180919
      /organism="Drosophila melanogaster"
      /mol_type="genomic DNA"
      /strain="Y: cn bw sp"
      /db_xref="taxon:7227"
      /chromosome="X"
      /map="10A2-10B2"
      /clone="BACR49A04 (D698) RPCI-98 49.A.4"
      /clone_lib="RPCI-98 (Roswell Park Cancer Institute
      Drosophila melanogaster BAC library, partial EcoRI in
      pBACE3.6"
      718..797
      /estimated_length=unknown
      1346..1425
      /estimated_length=unknown
      2042..2121

Query Match      31.4%; Score 967.8; DB 14; Length 180919;
Best Local Similarity 87.0%; Pred. No. Se-175;
Matches 1055; Conservative 0; Mismatches 82; Indels 76; Gaps 1;

QY 1697 CAAATACACACCCCAAGCGTTGGTGAACCAATTTTCGAGGCGCGAATGGCGACGTGCTT 1756
      |||||
Db 44625 CAGATACACACCCCAAGCGTTGGTGAACCAATTTTCGAGGCGCGAATGGCGACGTGCTT 44566

QY 1757 CGCTACGGCCGACGCGGATCGGGCAAAACGACACCACTGGCGCGGTGAGTTTAATGGAAA 1816
      |||||
Db 44565 CGCTACGGCCGACGCGGATCGGGCAAAACGACACCACTGGCGCGGTGAGTTTAATGGAAA 44506

QY 1817 GGTCCAGACTGCAAGAACGCGCATCTACGCCATGGCGGCCAAGATGCTTTGTGACCCCT 1876
      |||||
Db 44505 GGTCCAGACTGCAAGAACGCGCATCTACGCCATGGCGGCCAAGATGCTTTGTGACCCCT 44446

QY 1877 GAATATCCCGGTTTACC CGCCCATGAATCTAGTCTCTCGGCCAGTTCTTTGAGATTTA 1936
      |||||
Db 44445 GAATATCCCGGTTTACC CGCCCATGAATCTAGTCTCTCGGCCAGTTCTTTGAGATTTA 44386
```

```
QY 1937 CAGTGGCA----- 1944
Db 4385 CAGTGGCAAGGTGAGTGAGACAACATCAACTTGGCATGATTAATCATACTTAGCTGTATT 4326
QY 1945 -----AGTCTTCGATCTTCTGTGCGAAGCAAGCAAACTGCG 1980
Db 4325 AATCGATCGTTGTGTGTTGTCATAGGCTTCGATCTTCTGTCGCAAGCAAGCAAACTGC 44266
QY 1981 GCGTCTCGAGGATGTTAAACAGCAAGTGCAGGTGTCGGGACTCACCGAAGGTGTCG 2040
Db 4265 GCGTCTCGAGGATGTTAAACAGCAAGTGCAGGTGTCGGGACTCACCGAAGGTGTCG 44206
QY 2041 ATGGCGTCGAGGAGTACTGAAGCTCATCCAGCACGCAATGCTGCCGGAACATCCGCGC 2100
Db 4205 ATGGCGTCGAGGAGTACTGAAGCTCATCCAGCACGCAATGCTGCCGGAACATCCGCGC 44146
QY 2101 AGACGTGGCCAACTTCCAAATTCGTGCGCTTCGCGAGCGCTTTCAGATTTGCTGCGGC 2160
Db 4145 AGACGTGGCCAACTTCCAAATTCGTGCGCTTCGCGAGCGCTTTCAGATTTGCTGCGGC 44086
QY 2161 GCGAGGCTCGACGAAGATCCATGCAAGTTCCTGTTTCATCGATCGTGGCGCAATGAGC 2220
Db 4085 GCGAGGCTCGACGAAGATCCATGCAAGTTCCTGTTTCATCGATCGTGGCGCAATGAGC 44026
QY 2221 GGGCGTGGACACTTCTCGCGCGATCGGACGAGACGCTATGAGGGTCCGAGATTAAACA 2280
Db 4025 GGGCGTGGACACTTCTCGCGCGATCGGACGAGCGCTATGAGGGTCCGAGATTAAACA 43966
QY 2281 AATGCTCTGGCCCTCAAGGAGTGATTCGTGCGTTGGGCAACAGTCGCGCCCACTTGC 2340
Db 4365 AATGCTCTGGCCCTCAAGGAGTGATTCGTGCGTTGGGCAACAGTCGCGCCCACTTGC 43906
QY 2341 CCTTCGCTCTCAAACTCCACCAGTGTGCGGACTCTGTTTCATTCGCGCAGAGGCA 2400
Db 4305 CCTTCGCTCTCAAACTCCACCAGTGTGCGGACTCTGTTTCATTCGCGCAGAGGCA 43846
QY 2401 AGACGTGATGATAGCCATGATCTCGCGGACTTAGCTCTCGGAGCACACGCTCAACA 2460
Db 43845 AGACGTGATGATAGCCATGATCTCGCGGACTTAGCTCTCGGAGCACACGCTCAACA 43786
QY 2461 CGTTCGCTATGCGGATCGTCAAGGAGTGTGTGTCAGGATATCGTGAAGTTTGGC 2520
Db 43785 CGTTCGCTATGCGGATCGTCAAGGAGTGTGTGTCAGGATATCGTGAAGTTTGGC 43726
QY 2521 CTGGCGGACACCGAGCCATCGAGTACGCGAGCAGGAGGAGGAGGAGCTCAACA 2580
Db 43725 CTGGCGGACACCGAGCCATCGAGTACGCGAGCAGGAGGAGGAGGAGGAGCTCAACA 43666
QY 2581 TGGTGATCCGCACTCGCATCGATCGATCCCAATTCGATGCGACCGCCAGCGAGTCGA 2640
Db 43665 TGGTGATCCGCACTCGCATCGATCGATCCCAATTCGATGCGACCGCCAGCGAGTCGA 43606
QY 2641 ATAATCAGCGTCTCGGCTCTCATCTCGGGGCGGTCAATTCACAAATAATAATA 2700
Db 43605 ATAATCAGCGTCTCGGCTCTCATCTCGGGGCGGTCAATTCACAAATAATAATA 43546
QY 2701 ACAACAAACGAAACGCGCAACATCGGACCTTGGCCATCGTGGTCTGAGCG 2760
Db 43545 ACAACAAACGAAACGCGCAACATCGGACCTTGGCCATCGTGGTCTGAGCG 43486
QY 2761 AACACGAGATGTCGACGAGCTGATTTGTGACGACCGAGCCATCGACGACCTGACGAG 2820
Db 43485 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 43426
QY 2821 CGGAGGAGATGGT 2833
Db 43425 NNNNNNNNNNNAT 43413
```

RESULT 11  
BT001278  
LOCUS  
DEFINITION  
Drosophila melanogaster AT07729 full insert cDNA.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BT001278  
BT001278.1 GI:25009717  
FLI\_CDNA.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 2518)

REFERENCE  
AUTHORS

Scapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,  
Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,  
George, R., Gonzalez, M., Guarini, H., Kronmiller, B., Li, P., Liao, G.,  
Miranda, A., Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,  
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Rubin, G.M.  
and Ceiniker, S.

TITLE  
JOURNAL

Direct Submission  
Submitted (15-NOV-2002) Berkeley Drosophila Genome Project,  
Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, USA

COMMENT

Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to  
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
Science 2000). The sequence has been subjected to integrity checks  
for sequence accuracy, presence of a polyA tail and contiguity  
within 100 kb in the genome. Thus we believe the sequence to  
reflect accurately this particular cDNA clone. However, there are  
artifacts associated with the generation of cDNA clones that may  
have not been detected in our initial analyses such as internal  
priming, priming from contaminating genomic DNA, retained introns  
due to reverse transcription of unspliced precursor RNAs, and  
reverse transcriptase errors that result in single base changes.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our Web site  
(http://fruitfly.berkeley.edu) or send email to  
cdna@fruitfly.berkeley.edu.

FEATURES  
source

1..2518  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"

gene

1..2518  
/gene="CG12192"  
/db\_xref="FLYBASE:FBgn0034827"

CDS

168..2357  
/gene="CG12192"  
/notes="Longest ORF"  
/codon\_start=1  
/product="AT07729p"  
/protein\_id="AA071034.1"  
/db\_xref="GI:25009718"

translations="MDRIKIGEOLLQFQSRVHVQVCTAKNLERDSITGEMTEGTVR  
KGEVPLSTLIGINHHI FAENOPPI SKPPSLLPKRGSPTGPKSHTGNPVAERM  
VQDRSKIATRTDPAVAGNGHELGRAHTPRVPASRPDRDASPTQGRKSGGNNNQ  
RPSVVRNVRMGEQKRRARQEQKADLRNNPGRNPNWVSVMLRQYRLTLIF  
SPURCLDPNGTQQITVCVRKRPMSKEENKLDIITVPSDSLIVHELRLKVDLT  
KEFLHKKFRFDYTFDECSNLYDHTARPLIRTFEGGNATCFAYGQTSGKHTMG  
GEPFGVQDCTGTYAARDVPEVRPEYROMAKITCSFPEIYGVKVDLLPNK  
PMLRVLEDARQVWVGLTETMPVTQKEDVLRLTEHSGKERTSGTQSANAKSSSHAVP  
QIALHFPDSNGHGKCSFVLDLAGNKGADTQSDADROTREGAINKSLALKRCIRAL  
SROSHLPFGSKLTQVLKDSFVGKKKNTMIAMI SPSMSVENTILURYADRVKE  
LTAKEDHLQSGVEGDEKSPDLNEEPEMMADEGDEPEDNQHLTSSBEASY  
NNMSYDMSFNHTLNLIGSRNVDIQEVAQHALLVENLETIAHFNQKTDKTEQYT  
QNSESALMKLLAMVNTRDVTHTNYNTQKLLKEENQAYKKGELDESE"

ORIGIN

Query Match 18.5%; Score 568.6; DB 2; Length 2518;  
Best Local Similarity 65.8%; Pred. No. 1.6e-98;  
Matches 877; Conservative 0; Mismatches 444; Indels 12; Gaps 3;

```
1188 ACATTGCCCAATCCACGCGCGCAGTGTGTTCCGCGGCACAAAGAGTCCGCACT 1247
Db      AGAGCGCACACCCCTCTGTGTCCAGCCAGCTAGCGCGGACCGCGGTGTCAGCCCA 611
1248 GCGGCCCAACCCAGGAGCTGGCGGCGCTAGTACCCGCGCATCGCACGCAATTGAAGAG 1307
Db      GGTGCGCAAGTCAAGGGTGGCAACAATAGCAACCAACGCTTTTCCAGTGTAGTGGGAG 671
1308 GTGAGGAGCTGAAGAGAAATCCGAGAGACGACGCGCCGCAAGCGCGAGATGAAGAG 1367
Db      GTGAATTCGATGAAGAGCAGAGAGAGAGAGAGGAGGCTGCCAGGCGGAACAGCTCCAG 731
1368 GAGAAGGTGGCGTGTATGAACACAGGATCCGGGCAATCCAAACTGGGAGCGCGCAATG 1427
Db      GAGAAGGATGCACTGGTGCATTAATCCGGGAAATCCCTAACTGGGAGGTGTCGGTGATG 791
1428 ATACGCGAATATCAGAGCAGCTGGAATTTGTGCCGCT-----GTCGATGCCAGGCG 1481
Db      CTGCGCAATACCGCTCCACCTTGATCTTTCTCCACTTCGATCGCTGGATCCCAATGGA 851
1482 GTCGATGACCATCAGATCAGATGTGCGGTGGCGCAAGCGTCCCATTAAGCCGAAGGAGTC 1541
Db      GGCACCTGTTTCAGCAAAATACGGTGTGTGTGCGAAAAACGACCCATGATGTCGAAGGAG 911
1542 AATCGCAAGGAGATCGATGTCATTTCCGTGCCGCGCAAGGACATGCTCATCTGTCAGAG 1601
Db      AACTCCAAGAACCTGGACATCATCAGATTCCAGTGCAGGACAGCTGATCGTCCACGAG 971
1602 CCGCGCAGCAAGGTGCACCTCACCAGTTCTCTGGAGAACACAAAGTTTCGTTCCGACTAC 1661
Db      TTTGCGCTCAAGGTGATCTCACCAGTTCTCTGGAGAACACAAAGTTTCGTTCCGACTAC 1031
1662 GCCTTCAACGACACGTGCGACAAATGCGATGATACAAATACACAGCCAAAGCGTTGGTG 1721
Db      AGCTTTCGACAGGAGTGTCCAAATGCGTGGTCTACGATCAGTGTGCTCGTTGATC 1091
1722 AAAACCAATTTTCAGCGCGGAGTGGCAGCTGCTTCGCTAGCGCCAGAGCGGATCGGCG 1781
Db      AGAACCATGTTGAGGGCGGCAACGCGACTTGTGTTGCTTACGGAACAACTGGCAGCGGA 1151
1782 AAAACGACACCATGCGCGGTGAGTTTAATGAAGGTGCAGACTGCAAGACGCGATC 1841
Db      AAAACGACACCATGCGCGGAGAAATCTTCGAAAGGTTTCAGATTCGCGTACCGGATC 1211
1842 TAGCGCATGGCGCAAGGATGTCTTTGTGACCCCTGAATATGCCGCTTACCGCGCCATG 1901
Db      TAGCGCATGGCAGCTCGCGATGTCTTCGAGGAGGTATCGCGCCGCGAGTACCGGCAATG 1271
1902 AATCTAGTGTCTCGGCCAGTTTCTTTGAGATTACAGTGGCAAGTCTTCGATCTTCTG 1961
Db      GGTGCCAAGATTACGTGCGAGCTTCTTCGAAATCTATGGCAACCAAGGTGTTCGATCTTGT 1331
1962 T---CCGACACGAGAACTGCGCGTCTCGAGGAGTGTAAACAGCAAGTGCAGTGGTG 2018
Db      CTACCCACACGCCCATGCTGCGGGTCTTAGAGGATGCCAGCAGCAGGTCTGTTGGTG 1391
2019 GGAATCACGAGAAGTGTGATGGCGTCCGAGGAGGATCTGAAGCTCATCCAGCACGCG 2078
Db      GGCCTAACCGAGATGCGGTGACCAAGTGGAGGATGCTCTGAGACTGATTGAGCACGCG 1451
2079 AATGTCGCCGAAATCCGGCCAGAGCTGCGCAACTCAATTCGTGCGGTTTCGACGCG 2138
Db      AGCAAGAGCGCACTTCGCGCCAAACATCGCGGAAACGCGCAAGTCACTCGCTTCCACGCG 1511
2139 GTTTTTCAGATTGTCTGCGCGCGCAGGCTCGACGAAGNTCCATGGCAAGTTCCTCGTC 2198
Db      GTCTTTCAATAGCACTCACTTTCCGATTCCTGGGCGCCACACGCGCAAGTGTCTCTTT 1571
2199 ATCGATCTGGCGGCAATGAGCGGGCGGTGGACACTTCTCGCGCCGATCGGACAGCGGT 2258
Db      GTGGACTTGGCGGCAATGAACGCGGGGCGGATACGCAATCGCGGATCGTCAAACTGCG 1631
2259 ATGAGGGTGGCGAGATTAACAAAATCGCTGTGGCCCTCAAGGAGTGCATTTGTCGCTG 2318
```

```
Db      ATCGAGGAGCCGAGATCAATAAATCTCTGCTGGCCCTCAAGGAGTGCATTTGAGCCCTC 1691
QY      GGCMAAACAGTTCGCGCCACCTTCCGCTTGTCTCAAACTCACCAGGTGCTGCGCGAC 2378
Db      AGCGCGCATCGAGTCACCTTCCCTTCCGCTCCAGTTCAGTTCACCAAGTGTGCGCGAC 1751
QY      TCGTTCAITTTGGCG---AGAAGAGCAAGAGCTGTCATGATAGCCATGATCTCGCGGACCT 2435
Db      TCTTTGTGCGCGGCAAGAAACAGACCTGTCATGATTGCCATGATATCGCCATCCATG 1811
QY      AGCTCTCGGACACACGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGAGCTGGTG 2495
Db      AGCTGCTGGAGAAATACGCTCAACACTCTACGTTACGACAGACAGGGTTAAGGAGCTCATA 1871
QY      GTCAAGGATATCG 2508
Db      GCCAGGAGAGCG 1884

RESULT 12
CQ577533
LOCUS      CQ577533      2190 bp      DNA      linear      PAT 02-FEB-2004
DEFINITION Sequence 5291 from Patent WO0171042.
ACCESSION CQ577533
VERSION   CQ577533.1 GI:41640335
KEYWORDS
SOURCE    Drosophila sp.
ORGANISM  Drosophila sp.
           Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
           Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
           Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS   Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
TITLE     Detection kits, such as nucleic acid arrays, for detecting the
           expression of 10,000 or more Drosophila genes and uses thereof
           Patent: WO 0171042-A 5291 27-SEP-2001;
           PE Corporation (NY) (US)
FEATURES
           Location/Qualifiers
           source
             1..2190
               /organism="Drosophila sp."
               /mol_type="unassigned DNA"
               /db_xref="taxon:7242"

ORIGIN

Query Match      18.3%; Score 564.6; DB 6; Length 2190;
Best Local Similarity 68.1%; Pred. No. 9.5e-98;
Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;

QY      1296 GCATTGAAGAGGTGGAGCGACTGAAGGAGAAATCGGAGAGCGACGCGCCGACAGGCC 1355
Db      493 GTAGTGAAGGAGGTGAATCGCATGAAGGAGCAGAGGAGAGCGAAGGGCTCGCCAGCGG 552
QY      1356 GAGATGAAGGAGGAGAGGTGGCGCTGATGAACCGAGATCCGGGCAATCCAACTGGGAG 1415
Db      553 GAACAGCTCCAGGAGAGGATGCACTGCGTCGCAATAATCCGGGGAATCCCACTGGGAG 612
QY      1416 ACGGCGAAATGATACGCGAATATCAGAGCAGCTGGAATTTGTCCGCT-----GCTC 1469
Db      613 GTGTCGTGATGTGCGGCAATACCGTCCACCTTGATCTTTTCTCCACTTCGATGCGCTG 672
QY      1470 GATGCCAGGCGCTGCGATGACCATCAGATCAGTGTGCGTGCAGAGCGTCCCAATTAGC 1529
Db      673 GATCCCATGAGGAGCACTGTTTCAGCAATTAAGGTGTGTGTGCGAAGAACCCATGATG 732
QY      1530 CGCAAGGAGGTCAATCGCAAGAGATCGATGTCATTTTCGGTCCGCGCAAGGACATGCTC 1589
Db      733 CGCAAGGAGGAGAACTCCCAAGAACCTGGACATCATCAGTTCCAGTTCGCGACGACGCTG 792
QY      1590 ATCGTGCAGAGCGCGCAGCAAGGTGACCTCACCAGTTTCCTGGAGAACCAAGTTT 1649
Db      793 ATCGTCCATGATGTGGCGCTCAAGGTGATCTCACCAGATTCTCTGGAGCAACCAAAATTC 852
```

1650 CCCTTCGACTACGCTTCAACGACACGCTGCGCAATGCCATGCTATACAAATACACAGCC 1709  
853 CTTTTGCGACTACAGCTTCGACGAGAGTCTCCATGGCTGTCTACGATCACACTGCT 912  
1710 AAGCGGTGGTGAACCACTTTTCAGGGCGGAATGGCGAGCTGCTTCCCTTACGGCCAG 1769  
913 COTCGCTTGTATCAGAACCACTGTCAGGGCGCAATGCCACTGTGTTGCTTACCGACAA 972  
1770 ACGGATCGGGCAAAACGACACCATGGCGGTGAGTTTAATGGAAGAGTGCAGGACGTC 1829  
973 ACTGCGAGGGAAGAAACGACACCATGGCGGGAATTTCTCGAAGAGTTTCAGGATGTC 1032  
1830 AAGAACCGCATCTACGCCATGCGCGCAAGGATGCTTTGTGACCTGAAATATGCGCGT 1889  
1033 GGTACCGGGATCTACGCCATGCGCAGCTCGCGATGCTTCGAGGAGTATCGCGCCGAG 1092  
1890 TACCGCGGCATGAATCTAGTGTCTCGGCGAGTTCTTTTGAGATTTACAGTGGCAAGTTC 1949  
1093 TACCGGCAAAATGGGTGCCAAGATTACGTGACGCTTTCTTCGAAATCTATGGCACCAGGTG 1152  
1950 TTGATC--TTCTGTCGACAAGCAGAAACCTGCGCGCTCTGGAGATGGTAAACAGCAA 2006  
1153 TTGATCTCTGCTACCAACAGCCCATGCTGCGGGTCTTAGAGGATGCCAGGACGAG 1212  
2007 GTGCAAGTGTGGGACTCACCGAAGAGTGTGTCGATGCGGTGCGAGGAGTACTGAAGTTC 2066  
1213 GTGCGGTGTGGGCTTAACGAGATGCGGTGACCAAGTGGAGGATGCTTGAGACTG 1272  
2067 ATCAGCAGGGAATGCTGCGGCAACATCCGCGCAGAGCTGGGCAATCTCAATTTGTCG 2126  
1273 ATTGAGCAGCGCAGCAAGAGCGCACTTCCGCGCAACATCGGCGAAGCGCAAGTTCATCG 1332  
2127 GTTTCGACGCGCTTTTCCAGATTGTCTGCGCGCAGGCTGACCAAGATCCATGGC 2186  
2333 GTTCCACGCGCTTTTCAATAGACTCTCACTTTCCGANTCTGGGGGCCACACGGC 1392  
2187 AAGTCTCTGTTTCATGATCTGCGGCAATGAGCGGGCGTGGACACTTCTCGGCGGAT 2246  
1393 AAGTCTCTTTGTGACTTGGCGCAATGAACGCGGGCGGATACGCAATCCGCGAT 1452  
2247 CGGCAAGCGGTATGGAGGGTCCGAGATTAAACAAATGCTGCTGGCCCTCAAGAGATGC 2306  
1453 GTCAAACTCGCATCGAGGAGCGGAGATCAATAAATCTCTGCTGGCCCTCAAGAGATGC 1512  
2307 ATTGCTGTTGGCAACAGTTCGCCCACTTGGCTTCCGTGCTCCCAACTCACCCAG 2366  
1513 ATTGAGCCCTCAGCGCGAGTGCAGTCACTTCCCTTCCGTGCTCCCAAGTTCACCCAA 1572  
2367 GTGCTGCGGACTCGTTCAITGGCG--AGAGAGCAAGAGTGCATGATAGCCATGATC 2423  
1573 GTGCTGCGGACTCTTTGTGCGGGCAAGAGAACAGACTGCTGATGATGCTCATGATA 1632  
2424 TCGCGGAGACTTATGCTCTGCGAGACACGCTCAACAGCTGCGCTATGCGGATGCTGTC 2483  
1633 TCGCCATCCATGAGTGTGCGGAGAAATACGCTCAACACTCTACGCTTACGAGACAGGTT 1692  
2484 AAGAGCTGGTTCAGGATATCG 2508  
1693 AAGGAGCTCATAGCAAGGAAGACG 1717

RESULT 13  
CQ577532  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CQ577532  
Sequence 5290 from Patent WO0171042.  
CQ577532  
CQ577532.1 GI:41640334  
Drosophila sp.  
Drosophila sp.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

linear PAT 02-FEB-2004

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. 4190  
/organism="Drosophila sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:7242"

Query Match 18.3%; Score 564.6; DB 6; Length 4190;  
Best Local Similarity 68.1%; Pred No. 9.8e-98;  
Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;

Qy 1296 GCATTGAAGAGGTGGAGCGACTGAAGAGAAATCGCAGAGACGACGCGCCGACAGGCC 1355  
Db 1493 GTAGTGAAGGAGGTGAATCGCATGAAGGAGGAGGAGCGAAGGGCTCCGAGGCG 1552  
Qy 1356 GAGATGAAGGAGGAGGAGGTGGCGCTGATGAACAGGATCCGGCAATCCAACTGGGAG 1415  
Db 1553 GAAACAGCTCCAGGAGAGGATGCATCGCTGCGCAATAATCGGGGAATCCCACTGGGAG 1612  
Qy 1416 ACGCGCAAAATGATACCGGAATATCAGAGCAGCTGGAATTTGTGCCGCT-----GCTC 1469  
Db 1613 GTGTGCGTGTGCTGCGGCCAATACCGTCCACCTTGATCTTTTCTCCACTTCGATGCTG 1672  
Qy 1470 GATGCCAGGCGCTGATGACCATCAGATCAGATGTCGTCGCAAGCGTCCCAATTAGC 1529  
Db 1673 GATCCCAATGAGGCACTGTTAGCAATTTACGGTGTGTGTCGAAACGACCATGAT 1732  
Qy 1530 CGAAGAGGTCAATCCAAAGGAGATCGATGTCAATTTGGTGGCGCGCAAGGACATGCTC 1589  
Db 1733 CGCAAGGAGGAGAACTCCAAGAACCTGAGCATCATCAGTTCCAGTTCGCGACAGCCTG 1792  
Qy 1590 ATCGTGACAGGCGCGCAGCAAGTGCACCTCAAGTTTCTGGAGAGACCAACAAGTTT 1649  
Db 1793 ATCGTCCATGATGTCGCCCTCAAGTGGATCTCACCAGTTCTCTGGAGACCAACAATTC 1852  
Qy 1650 CGTTTCGACTACGCTTCAACGACACATGCGACAAATGCCATGGTATACAAATACACAGCC 1709  
Db 1853 CGTTTCGACTACAGTTTCGACGAGGAGTGTCCAAATGGCTGTCTACGATCACACTGCT 1912  
Qy 1710 AAGCGGTGGTGAACCACTTTTCAGGGCGGAAATGGCGAGCTGCTTCCGCTACGGCCAG 1769  
Db 1913 CGTCGCTGTGATCAGAACCACTGTTTCAGGGCGGCAATGCCACTTGTCTTCGCTTACGACAA 1972  
Qy 1770 ACGGATCGGGCAAAACGACACCATGGCGGTGAGTTTAATGGAAGAGTGCAGGACTGC 1829  
Db 1973 ACTGCGAGGAAACGACACCATGGGCGGAGAAATCTTCGAAAGGTTTCAAGATTGC 2032  
Qy 1830 AAGAACCGCATCTACGCGCATGGCGGCAAGGATGCTTTGTGACCTTGAATATACCGCGT 1889  
Db 2033 GGTACCGGGATCTACGCCATGGCAGCTCGCGATGCTTCGAGGAGGATTCGCGCCCGAG 2092  
Qy 1890 TACCGCGCATGAATCTAGTGTCTCGGCCAGTTTCTTTTGAGATTTACAGTGGCAAGTTC 1949  
Db 2093 TACCGGCAAAATGGGTGCCAAGATTACGTGACGCTTCTTCGAAATCTATGGCACCAGGTG 2152  
Qy 1950 TTGATC--TTCTGTCGACAAGCAGAACTCGCGTCTCGGAGGATGGTAAACAGCAA 2006  
Db 2153 TTGATCTCTGCTACCAACAGCCCATGCTCGGGTCTTAGAGGATGCCAGGACGAG 2212  
Qy 2007 GTGCAAGTGTGGGACTCACCGAAGAGTGTGTCGATGCGCTGCGAGGAGTACTGAAGTTC 2066  
Db 2213 GTGCTGCTGGGCTTAACGAGATGCGCGGTGACCAAAAGTGGAGGATGCTCTGAGACTG 2272  
Qy 2067 ATCAGCAGGCAATGCTGCCGAAACATCCGCGCAGAGCTGCGGCAACTCCAATTCGTCG 2126  
Db 2273 ATTGAGCAGCGCAGCAAGAGCGCACTTCCGCGCAACATCGGCGCAACGCGCAAGTTCATCG 2332

```
QY 2127 CGTTCGACGCGGTTTTCCAGATTGTGCTGGCGCCGAGGCTCGACGAAGATCCATGCGC 2186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2333 CGTTCGACGCGGCTTTTCAATAGCACTCCACTTTCCGGATTCTCTGGGGCCACACGCGC 2392
QY AAGTCTCTGTTTCATCATCTGGCGGCAATGAGCGGGCGTGACACTTCTCTGGCGCGAT 2246
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2393 AAGTGTCTCTTGTGACTTGGCGGCATGACGCGGGCGGATACGCAATCCGCCGAT 2452
QY CGGCACAGCGGTATGAGGGTGCAGAGATTAACAATCGTCTGCGCCCTCAAGGAGTGC 2306
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2453 CGTCAAACTCGCATCGAGGAGCGGAGATCAATAATCTCTGCTGGCCCTCAAGGAGTGC 2512
QY ATTCTGCTGTTGGGCAACAGTGGCCCACTTGCCTTCCGTGTCTCAAACTCACCCAG 2366
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2513 ATTCTGAGCCCTCAGCGCGAGTGCAGTCACTTCCCTTCCGTGGTCCAAGTTGACCCAA 2572
QY GTGCTCGCGACTCGTTTCATTGSCG--AGAAGAGCAAGACGTGCGATAGCATGATC 2423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2573 GTGCTCGCGACTCTTTTGTGGCGGCAAGAGAACAGACCTGCATGATTCATGATA 2632
QY TCGCCGGGACTTAGTCTCTCGGAGCACACGCTCAACACGCTCGCTATCGCGATCGTGC 2483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2633 TCGCCATCCATGAGCTGCGTGAGGATACGCTCAACACTCTAGTTACGACAGCGGTT 2692
QY AAGAGCTGCTGCTCAAGATATCG 2508
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2693 AAGGAGCTCATAGCCCAAGGAGACG 2717
```

RESULT 14  
AC006245/c

## LOCUS

## DEFINITION

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## COMMENT

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## COMMENT

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## COMMENT

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## COMMENT

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## COMMENT

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## COMMENT

the following cutoffs: length >= 200 bases. P1 library location: 82-55.

\* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces

\* is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence

\* as soon as it is available and the accession number will be preserved.

\* 1 588: contig of 588 bp in length

\* 589 668: gap of unknown length

\* 669 1229: contig of 561 bp in length

\* 1230 1309: gap of unknown length

\* 1310 82269: contig of 80960 bp in length

\* 82270 82349: gap of unknown length

\* 82350 82884: contig of 535 bp in length

\* 82885 82964: gap of unknown length

\* 82965 83532: contig of 568 bp in length

\* 83533 83612: gap of unknown length

\* 83613 84419: contig of 807 bp in length.

\* Location/Qualifiers

\* 1. 84419

organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/strain="y; cn bw sp"

/db\_xref="taxon:7227"

/chromosome="2"

/map="59D1-59D2"

/clone="DS07831 (D467)"

/clone\_lib="P1 library, partial Sau3A in pAd10sacBI"

589. .668

/estimated\_length=unknown

1230. 1309

/estimated\_length=unknown

82270. .82349

/estimated\_length=unknown

82885. .82964

/estimated\_length=unknown

83533. .83612

/estimated\_length=unknown

ORIGIN

Query Match 18.3%; Score 564.6; DB 14; Length 84419;

Best Local Similarity 68.1%; Pred. No. 1.1e-97;

Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;

QY 1296 GCATTGAAGAGGTGGAGCGACTGAAGAGAGATCCGAGAACGCGCGCGACAGGCC 1355

Db 31336 GTAGTGAGGGAGGTGAATCGCATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 31277

QY 1356 GAGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1415

Db 31276 GAACAGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 31217

QY 1416 ACGCGGCAATGATACGCAATATCAGAGCAGCGTGGAAATTTGTGCCCTG-----GCTC 1469

Db 31216 GTGTGGTGATGCTGGGCAATACCGCTCCACCTTGATCTTTTCTCCACTTCGATGCTG 31157

QY 1470 GATGCCAGGCGCGTGGATGACATCAGATCAGTGTGCGTGGCGCAAGCGTCCCATAGC 1529

Db 31156 GATCCCAATGAGGCACTGTTTCAGCAATATACGGTGTGTGTCGCAAAACGACCCATGAT 31097

QY 1530 CGCAGGAGGTCAATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1589

Db 31096 CGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 31037

QY 1590 ATCTGTACGAGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1649

Db 31036 ATCTGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 30977

QY 1650 CGCTTCGACTACGCTTCAACGACGCGTGGCAATGCCATGGTATACAAATACACAGCC 1709

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 30976 CTTTTCAGTACACGTTTCGAGGAGGTGCTCCAAATGCGTGGTCTACGATCACACTGCT 30917  
 QY 1710 AAGCGTGTGGTGAACCAATTTTCAGGCGGGAATGCGACGTCTTCGCTCAGGCCAG 1769  
 Db 30916 CTTTTCAGTACAGAACCATGTTTCAGGCGGGAATGCGACGTCTTCGCTCAGGCCAG 30857  
 QY 1770 ACGGATCGGGCAAAACCCACACCATGCGGCGGTGATTTAATGGAAGAGTGCAGGACTGC 1829  
 Db 30856 ACTGCGAGCGGAAACCCACACCATGCGGCGGGAATTTCTTCGGAAGAGTTCAGGATTCG 30797  
 QY 1830 AAGACGCACTACGCTATGCGGCGGCAAGATGCTTTGTGACCTGAAATATGCGCGGT 1889  
 Db 30796 GGTACCGGGATCTACGCGATGCGACTGCGATGCTTCGAGGAGGTATCGCGCCGAG 30737  
 QY 1890 TACCGCGCATGAATCTAGTCTGCTCGGCCAGTTCTTTTGAGATTTACAGTGGCAAGTTC 1949  
 Db 30736 TACCGGCAATGGTGCCAAATGATGCTGAGCTTTCTTCGAAATCTATGGCACCAGGTG 30677  
 QY 1950 TTGATC---TTCTGTCCGACAGCAAACTGCGCGTCTCTGGAGGATGGTAAACAGCAA 2006  
 Db 30676 TTTGATCTTTGCTACCCCAACAGCCCATGCTGCGGGTCTTAGAGGATGTCAGGCGAGC 30617  
 QY 2007 GTGCAAGTGTGGGACTCACGAGAGTGTGATGCGCTGAGAGGATGCTGAGGCTC 2066  
 Db 30616 GTGCTGTGGTGGGCTTAACGAGATGCGCGTGACCAAGTGGAGGATGCTCTGAGACTG 30557  
 QY 2067 ATCCAGCAGCGAATGCTGCGCGCAATCTCGGCGGCGGAGGCTCGAGAGATCCATGGC 2186  
 Db 30556 ATTGAGCAGCGAGCAAGAGCGCACTTCCGCGCAAAATCGGCAAGCCCAATCATCG 30497  
 QY 2127 CGTTGCGACGCGTTTTCAGATTTGTGCTGCGGCGGCGGAGGCTCGAGAGATCCATGGC 2186  
 Db 30496 CGTTCCCAAGCGCTTTTCAAATAGCACTCCACTTTTCCGATTTCTGGGGCCACACGGC 30437  
 QY 2187 AAGTCTCTGATCATGCTGCGGCGCAATGAGCGGGCGTGGACACTTCTCTCGGCGCAT 2246  
 Db 30436 AAGTCTCTTTTGTGAGCTTGGCGGCAATGAGCGGGCGGATAGCGCAATCGCGCAT 30377  
 QY 2247 CGGAGAGCGCTATGAGGGTGGCGAGATTAAACAAATGCTGCTGGCGCTCAAGGATGC 2306  
 Db 30376 CGTCAATCTGATGAGGAGCGGAGATCAATAATCTCTGCTGGCGCTCAAGGATGC 30317  
 QY 2307 ATTCTGCTGTGGGCAACAGTCTGCGGCGCACTTGGCTGCTCCAACTCACCCAG 2366  
 Db 30316 ATTCTGAGCGCTCAGCGGAGTCTGAGTCACTTCCCTTCCGCTGCTCAAGTCAACCA 30257  
 QY 2367 GTGCTGCGGAGCTGTTTATTTGCGG---AGAAGAGCAAGCTGCTGATGATGATGATC 2423  
 Db 30256 GTGCTGCGGAGCTCTTTGCTGCGGCGCAAGAGCAAGACCTGCTGATGATGATGATG 30197  
 QY 2424 TCGCGGAGCTTAGTCTCTGCGGAGCAAGCTCAACAGCTGCTGCTGCTGCTGCTGCTGCT 2483  
 Db 30196 TCGGATCCATGAGCTGGTGGAGATAGCTTCAACTTACCTTACCTTACCTTACCTTACCT 30137  
 QY 2484 AAGAGCTGTGGTCAAGGATATCG 2508  
 Db 30136 AAGAGCTCATGACCAAGAGAGCG 30112

RESULT 15  
 AC099036/c 180361 bp DNA linear INV 08-NOV-2001  
 LOCUS Drosophila melanogaster, chromosome 2R, region 59C-59D, BAC clone  
 DEFINITION BACR34K16, complete sequence.  
 AC099036  
 AC099036  
 AC099036.1 GI:16798968  
 HTG.  
 Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 180361)

AUTHORS

Celniker, S.E., Adams, M.D., Krommiller, B., Tyler, D., Wan, K.H.,  
 Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C.,  
 Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A.,  
 Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M.,  
 Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D.,  
 Ferreira, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A.,  
 Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J.,  
 Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattel, B., Moshrefi, A.,  
 McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J.,  
 Pacleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B.,  
 Phouanenavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F.,  
 Stapleton, M., Strong, R., Svirskas, R., Tector, C., Williams, S.M.,  
 Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.

Sequencing of Drosophila chromosome 2R, region 59C-59D

Unpublished

2 (bases 1 to 180361)

TITLE

JOURNAL

REFERENCE

AUTHORS

Celniker, S.E., Adams, M.D., Krommiller, B., Tyler, D., Wan, K.H.,  
 Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C.,  
 Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A.,  
 Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M.,  
 Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D.,  
 Ferreira, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A.,  
 Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J.,  
 Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattel, B., Moshrefi, A.,  
 McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J.,  
 Pacleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B.,  
 Phouanenavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F.,  
 Stapleton, M., Strong, R., Svirskas, R., Tector, C., Williams, S.M.,  
 Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.

Direct Submission

TITLE

JOURNAL

REFERENCE

AUTHORS

COMMENT

Submitted (08-NOV-2001) Berkeley Drosophila Genome Project, MS  
 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,  
 Berkeley, CA 94720, US  
 Sequence submitted by:  
 Berkeley Drosophila Genome Project  
 Lawrence Berkeley National Laboratory, MS 64-121  
 Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome  
 shotgun and from subclones of this BAC and its neighboring clones.  
 For further information about this sequence, including its location  
 and relationship to other sequences, please visit our sequence  
 archive web site (<http://www.fruitfly.org/sequence/>) or send email  
 to [bdgp@fruitfly.berkeley.edu](mailto:bdgp@fruitfly.berkeley.edu).

FEATURES

source

1..180361  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /strain="y; cn bw sp"  
 /db\_xref="taxon:7227"  
 /chromosome="2R"  
 /map="59C-59D"  
 /clone="BACR34K16 (D1381)"  
 /clone\_lib="RPCI-98 (Roswell Park Cancer Institute  
 Drosophila melanogaster BAC library, partial EcoRI in  
 pBACE3.6)"

ORIGIN

Query Match 18.3%; Score 564.6; DB 2; Length 180361;  
 Best Local Similarity 68.1%; Pred. NO. 1.1e-97;  
 Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;  
 QY 1296 GCATTGAAGAGGTGGAGGAGTGAAGGAGAAATCGCAGAGGAGCGCGCCGACAGGCC 1355  
 Db 133683 GTAGTAGGGAGGTGAATCGCATGAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGG 133624  
 QY 1356 GAGATGAAGGAGGAGGAGGAGTGGCGCTGATGAACACAGGATCCGGGCAATCCAACTGGAG 1415  
 Db 133623 GAACAGCTCCAGGAGGAGGATGCACTGCGTCCGCAATAATCCGGGGAATCCCACTGGAG 133564  
 QY 1416 ACGCGCAATGATAGCGAATATCAGAGCAGCTGGAATTTGTGGCT-----GCTC 1469  
 Db 133563 GTGTCGGTGTATGCTGCGCAATACCGCTCCACCTTGATCTTTTCCACTTCGATCGCTG 133504



Search completed: May 12, 2006, 21:01:43  
Job time : 15365 secs

QY 1470 GATGGCCAGCCGTCGATGACCATCAGATCAGAGTGTGGTGGCGAAGCGTCCATTAGC 1529  
Db 133503 GATCCCAAATGGAGGCACTGTTCAGCAAAATACCGGTGTGTGTGCGAAAACGCCCATGAGT 133444  
QY 1530 CGCAAGGAGGTCAATCGCAAGGAGATCGATGTCTATTTCCGTGCGCGCAAGGACATGCTC 1589  
Db 133443 CGCAAGGAGGAGAACTCCAGAAGACCTTGGACATCATCAGTTCCTGAGTCCGAGACGCTTG 133384  
QY 1590 ATCGTGCAAGCGCGCGCAGCAAGGTGACCTCACCAAGTTCCTGGAGAACCAACAAGTTT 1649  
Db 133383 ATCGTCCATGAGTTGGCGCTCAAGTGGATCTCACCAAGTTCCTGGAGCACCAAAATTC 133324  
QY 1650 CGCTTCGACTACGCTTCAACGACACGTGGGCAATGCCATGCTATACAAATACACAGCC 1709  
Db 133323 CGTTTCGACTACAGTTGACAGGAGTGTCTCAATGCGTGGTCTACGATCACACTGCT 133264  
QY 1710 AAGCCGTGTGTGAAAACCAATTTTCGAGGGCGGAATGGCGACGTGCTTCGCCTACGGCCAG 1769  
Db 133263 CGTCCGTTGATCAGAACCAATGTTTCGAGGGCGGCAATGCCACTTGTTCCTTACGGACAA 133204  
QY 1770 ACGGATTCGGGCAAAAACGACACCATGGGCGGTGAGTTTAAATGGAAGGTGCAGGACTGC 1829  
Db 133203 ACTGGCAGCGGAAAAACGACACCATGGGCGGAGAAATCTTCGGAAGGTTCAGGATTGC 133144  
QY 1830 AAGAACGGGATCTACGCCATGGCGGCAAGGATGTCTTTGTGACCTCGAATATGCCGCGT 1889  
Db 133143 GGTACCGGGATCTACGCCATGGCAGCTCGCGATGTCTTCGAGGAGGTATCGGCGCCGGAG 133084  
QY 1890 TACCGGCGCATGAATCTAGTCTGCTCGGCCAGTTCTTTGAGATTTACAGTGGCAGGTC 1949  
Db 133083 TACCGGCAATGGGTGCCAAGATTACGTGAGCTTCTTCGAAATCTATGGCACCAAGGTG 133024  
QY 1950 TTGATC---TTCTGTCCGACAGCAAGAACTGCGGCTCTTGAGGATGTTAAACAGCAA 2006  
Db 133023 TTGATCTCTTGTCTACCAACAAAGCCATGCTCGGGTCTTAGAGGATGCCAGGACAGCAG 132964  
QY 2007 GTGCAGGTGGTGGACTACCGGAGAGGTGTGATGGCGTTCGAGGAGGTACTGAAGTTC 2066  
Db 132963 GTCGTGTGTGGGCTTAAACGGAGATGCGGTGACCAAAAGTGGAGGATGTCTCGAGACTG 132904  
QY 2067 ATCCAGCAGCGCAATGCTCCGGAACATCGGCCAGACGTGGGCCAACTCCAAATTCGTG 2126  
Db 132903 ATTGAGCACGGCAGCAAGAGCGCACTTCGGCCAAACATCGCGCAACGCCAAGTCATCG 132844  
QY 2127 CGTTCGACGCGGTTTTCAGATTGTGTGTCGGCGCGCAGGGCTCGACGAAGATCCATGGC 2186  
Db 132843 CGTTCCACGCGCTCTTTCAAATAGCACTCCACTTTCCCGATTCCTGGGGCCCAACAGGC 132784  
QY 2187 AAGTTCTCGTTATCGATCTGGGGGCAATGAGCGGGGCGTGGACACTTCCTCGGCGCGAT 2246  
Db 132783 AAGTGTCTCTTTGTGACTTGGCGGCAATGAACGCGGGCGGATACGCAATCCGCGGAT 132724  
QY 2247 CGCGAGACGCGTATGAGGGTCCGAGATTACAAATCGCTGCTGGCCCTCAAGGAGTGC 2306  
Db 132723 CGTCAAACTCGCATCGAGGAGCCGAGATCAATAATCTCTGTGGCCCTCAAGGAGTGC 132664  
QY 2307 ATTCTGTGCTTGGGCAAAACAGTGGGCCACTTGGCCCTTCCTGTGTCTCCAAACTCACCCAG 2366  
Db 132663 ATTGAGCCCTCAGCGCGAGTCACTACCTTCCTTCCTGGTCCGAGTTGACCCAA 132604  
QY 2367 GTGCTGCGGACTCGTTTCAATGGCG---AGAAGAGCAAGCGTGATGATAGCCATGATC 2423  
Db 132603 GTGCTGCGGACTCTCTTTGTCGGCGCAAGAGAACAGACCTGCATGATGATTCATGATA 132544  
QY 2424 TCGCGGGGACTTAGCTCTCGGAGCACAGCTCAACAGCTGCGGTATGCGGATCGTGTTC 2483  
Db 132543 TCGCCATCCATGAGCTGCGTGAGGAAATCGCTCAACACTCTACGTTACGACAGAGGTT 132484  
QY 2484 AAGGAGCTGCTGCTCAAGGATATCG 2508  
Db 132483 AAGGAGCTCATAGCCDAGGAAGACG 132459

***This Page Blank (uspto)***

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 16:33:29 ; Search time 1766 Seconds  
(without alignments)  
11627.360 Million cell updates/sec

Title: US-10-840-060-142  
Perfect score: 3081  
Sequence: 1 aaacacaaaattgtgtgc.....gccgaatggcaagcgttagt 3081

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9991994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N Geneseq 21.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3081	100.0	3081	4	ABL02575
2	3081	100.0	3081	10	ADK11353
3	1547.8	50.2	8043	4	ABL02574
4	564.6	18.3	2190	4	ABL05367
5	564.6	18.3	4190	4	ABL05366
6	505.8	16.4	1970	4	ABL05833
7	505.8	16.4	3370	4	ABL05832
8	454.6	14.8	1292	6	AAD24080
9	454.6	14.8	1292	6	AAD24082
10	454.6	14.8	1292	12	ADO52609
11	454.6	14.8	1421	6	AAD24083
12	454.6	14.8	1421	10	ADE10088
13	454.6	14.8	1421	12	ADO52615
14	454.6	14.8	2172	6	AAD24086
15	454.6	14.8	2172	10	ADE10094
16	454.6	14.8	2172	12	ADO52621
17	453	14.7	2865	6	AAS94953
18	449.8	14.6	1175	6	AAD24081
19	449.8	14.6	1175	10	ADE10084

20	449.8	14.6	1175	12	ADO52611	Human mit
21	449.8	14.6	1304	6	AAD24084	Human mit
22	449.8	14.6	1304	10	ADE10090	DNA encod
23	449.8	14.6	1304	12	ADO52617	Human mit
24	449.8	14.6	2740	6	ABL65400	Lung canc
25	449.8	14.6	2740	6	ABL65848	Lung canc
26	449.8	14.6	2740	6	ABL64767	Lung canc
27	449.8	14.6	2740	6	ABK70308	Human lun
28	449.8	14.6	2740	6	ABK70303	Human lun
29	449.8	14.6	2740	6	AAD24087	Human mit
30	449.8	14.6	2740	9	ACA90156	Human cdn
31	449.8	14.6	2740	10	ADE10096	DNA encod
32	449.8	14.6	2740	10	AAD54025	Human col
33	449.8	14.6	2740	12	ADO52623	Human mit
34	449.8	14.6	2740	13	ADT90126	Human gen
35	449.8	14.6	2740	13	ADT90283	Human gen
36	449.8	14.6	2740	13	ADT89995	Human gen
37	449.8	14.6	2740	13	ADT89892	Human gen
38	449.8	14.6	2825	12	ADQ09301	Human KNS
39	449.8	14.6	2825	13	ADR25467	Breast ca
40	449.8	14.6	2825	14	ADX06781	Cyclin-de
41	449.8	14.6	2825	14	ADY16204	DNA encod
42	445.2	14.4	2892	12	ADQ22362	Human eof
43	441	14.3	2857	13	ADQ85075	Human tum
44	441	14.3	2857	13	ACN37373	Tumour-as
45	436.2	14.2	2741	3	AAF15699	Human pro

ALIGNMENTS

RESULT 1

ABL02575  
ID ABL02575 standard; cDNA; 3081 BP.

AC ABL02575;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 2207.

XX Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical; gene; ss.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB58472.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

XX Claim 1; SEQ ID NO 2207; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention



Db 1981 GCGTCTGAGGATGGTAAACAGCAAGTGCAGGTGTGGGACTCACCGAAGGTGGTCG 2040  
Qy 2041 ATGCGTGCAGAGGTAATGAGCTCATCAGACAGCAATGTGTCGCAACATCCGCGCC 2100  
Db 2041 ATGCGTGCAGAGGTAATGAGCTCATCAGACAGCAATGTGTCGCAACATCCGCGCC 2100  
Qy 2101 AGAGTCCGGCAACTCCAAATTCGTCGCTTCGACAGCGCTTTCCAGATTTGCTGGCG 2160  
Db 2101 AGAGTCCGGCAACTCCAAATTCGTCGCTTCGACAGCGCTTTCCAGATTTGCTGGCG 2160  
Qy 2161 CGCAGGCTCGACGAAGATCCATGGCAAGTTCTGTTTCATCGATCTCGCGGCAATGAGC 2220  
Db 2161 CGCAGGCTCGACGAAGATCCATGGCAAGTTCTGTTTCATCGATCTCGCGGCAATGAGC 2220  
Qy 2221 GGGCGTGGACACTTCTCGCGCATCGGCAGACGCTATGAGGGTGCAGATTAACA 2280  
Db 2221 GGGCGTGGACACTTCTCGCGCATCGGCAGACGCTATGAGGGTGCAGATTAACA 2280  
Qy 2281 AATCGCTGCTGGCCCTCAAGAGTGCATTCTGCTGTTGGGCAACAGTCCGGCCACTTGC 2340  
Db 2281 AATCGCTGCTGGCCCTCAAGAGTGCATTCTGCTGTTGGGCAACAGTCCGGCCACTTGC 2340  
Qy 2341 CTTTCCGTGTCTCAAACTCACCAAGTGTGCGCGACTCGTTTCATTGGCGAAGAGCA 2400  
Db 2341 CTTTCCGTGTCTCAAACTCACCAAGTGTGCGCGACTCGTTTCATTGGCGAAGAGCA 2400  
Qy 2401 AGACGTGATGATGACCATGATCTCGCGGCACTTACGCTCTCGAGACACAGCTCAACA 2460  
Db 2401 AGACGTGATGATGACCATGATCTCGCGGCACTTACGCTCTCGAGACACAGCTCAACA 2460  
Qy 2461 CGCTGGCTATGCGGATCGTGTCAAGAGCTGTGTCAAGGATATCGTCAAGTTTGGC 2520  
Db 2461 CGCTGGCTATGCGGATCGTGTCAAGAGCTGTGTCAAGGATATCGTCAAGTTTGGC 2520  
Qy 2521 CTGCGCGGACACCGAGCCCATCGAGATCACGACGACGAGGAGGAGGAGCTCAACA 2580  
Db 2521 CTGCGCGGACACCGAGCCCATCGAGATCACGACGACGAGGAGGAGGAGCTCAACA 2580  
Qy 2581 TGTGTATCGCACTCGCATCAGCTGCATCCCAATTCGATGACCGCGGCGAGCTCGA 2640  
Db 2581 TGTGTATCGCACTCGCATCAGCTGCATCCCAATTCGATGACCGCGGCGAGCTCGA 2640  
Qy 2641 ATAATCAGCTGTCTCGGCTCTCATCATCTCGGGGGGGTTCATTCACAAATTAATA 2700  
Db 2641 ATAATCAGCTGTCTCGGCTCTCATCATCTCGGGGGGGTTCATTCACAAATTAATA 2700  
Qy 2701 ACAACAAAGAACGAAACCGCGCAACATGACCTGGCCATGCTGAGTTGCTGAGCG 2760  
Db 2701 ACAACAAAGAACGAAACCGCGCAACATGACCTGGCCATGCTGAGTTGCTGAGCG 2760  
Qy 2761 AACACAGATGTCGACGAGTGAATGTGACGACGAGGCAATCGACGACCTGACGAGA 2820  
Db 2761 AACACAGATGTCGACGAGTGAATGTGACGACGAGGCAATCGACGACCTGACGAGA 2820  
Qy 2821 CGGAGAGATGTGTGGAGTATCATCGACCGTTAATGCGACACTGGAGACCTTCTCG 2880  
Db 2821 CGGAGAGATGTGTGGAGTATCATCGACCGTTAATGCGACACTGGAGACCTTCTCG 2880  
Qy 2881 CCGAGTCCGAAGGGCTGTAAATCTGACCAACTATGTGACTACGACGAGGACTCGTACT 2940  
Db 2881 CCGAGTCCGAAGGGCTGTAAATCTGACCAACTATGTGACTACGACGAGGACTCGTACT 2940  
Qy 2941 GCAACCGGGCGAGTGCATGTTCTTCGAGTGTGGACATCGCCATCCAGTCCGCGACA 3000  
Db 2941 GCAACCGGGCGAGTGCATGTTCTTCGAGTGTGGACATCGCCATCCAGTCCGCGACA 3000  
Qy 3001 TGAATGCCGATATCGGCGCAAGTTGGCCAGAGAGATGCTGTGCTGAGCTTCAATT 3060  
Db 3001 TGAATGCCGATATCGGCGCAAGTTGGCCAGAGAGATGCTGTGCTGAGCTTCAATT 3060  
Qy 3061 CGCCGAATGCGAGCGGTAGT 3081  
Db 3061 CGCCGAATGCGAGCGGTAGT 3081

## RESULT 2

ADK11353

ID ADK11353 standard; DNA; 3081 BP.

XX

AC ADK11353;

XX

DT 06-MAY-2004 (first entry)

XX

Drosophila kinesin-like protein KIF2 homolog gene.

XX

ds; gene; cytostatic; cardiovascular; immunosuppressive; nephrotropic; antineumatic; antiarthritic; dermatological; antiporiatic; antiinflammatory; fungicide; gene therapy; Drosophila; diagnosis; cardiovascular disorder; autoimmune disease; glomerulonephritis; rheumatoid arthritis; dermatological disorder; psoriasis; inflammatory disorder; malaria; emphysema; alopecia.

XX

Drosophila melanogaster.

XX

W02003040301-A2.

XX

15-MAY-2003.

XX

23-OCT-2002; 2002WO-GB004780.

XX

05-NOV-2001; 2001GB-00026506.

XX

27-NOV-2001; 2001GB-00028384.

PR

11-FEB-2002; 2002GB-00003185.

XX

(CYCL-) CYCLACEL LTD.

XX

Deak P, Frenz L, Glover D, Midgley C;

XX

WPI; 2003-441540/41.

XX

P-PSDB; ADK11354.

XX

New Drosophila polypeptides and polynucleotides, useful for diagnosing,

XX

preventing and/or treating disorders, such as cancer, glomerulonephritis, rheumatoid arthritis, psoriasis, malaria, emphysema and alopecia.

XX

Example 9; Page 129-130; 265pp; English.

XX

The invention relates to novel Drosophila species DNA sequences and their encoded proteins with their corresponding human homologues. The proteins or their encoding polynucleotides are useful in a method of prevention, treatment or diagnosis of a disease in an individual, and used to identify a substance capable of binding to the polypeptide or modulating the function of the polypeptide comprising incubating the polypeptide with a candidate substance and determining whether the substance binds to the polypeptide. The compositions are administered to an individual in need of such treatment. The method of diagnosis, in which the presence or absence of a polynucleotide is detected in a biological sample, comprises brining the biological sample containing the nucleic acid such as DNA or RNA into contact with a probe comprising a fragment of at least 15 nucleotides of the polynucleotide, and detecting any duplex formed between the probe and nucleic acid in the sample. The method also comprises providing an antibody capable of binding to the polypeptide, incubating a biological sample with the antibody to allow the formation of an antibody-antigen complex, and determining whether antibody-antigen complex comprising the antibody is formed. The disease comprises a proliferative disease such as cancer. The antibody or identified substance is also useful in inhibiting the function of a polypeptide and/or regulating a cell division cycle function. The diseases also include cardiovascular disorders, autoimmune diseases such as glomerulonephritis and rheumatoid arthritis, and dermatological disorders such as psoriasis, inflammatory, fungal, and parasitic disorders such as malaria, emphysema and alopecia. This sequence represents one of the Drosophila genes of the invention.

XX

Sequence 3081 BP; 803 A; 817 C; 868 G; 593 T; 0 U; 0 Other;

XX

SQ

Query Match		100.0%	Score 3081;	DB 10;	Length 3081;		
Best Local Similarity		100.0%	Pred. No. 0;				
Matched 3081;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	AAACTAAAAAATGTTGCTGCTGACATCTGGTGGCTTGCAGAACTATTTCTAGCAGATTTT	60				
DB	1	AAACTAAAAAATGTTGTTCTGCTGACATCTGGTGGCTTGCAGAACTATTTCTAGCAGATTTT	60				
QY	61	GTGATATTTCTGTTGTGATCGGTGCGATAAATCGCCAGTTTTTTTTTTTAAATGAAAGTGT	120				
DB	61	GTGATATTTCTGTTGTGATCGGTGCGATAAATCGCCAGTTTTTTTTTTTAAATGAAAGTGT	120				
QY	121	AACACATTTGACGGTTGGAGATAGCAGAGAAAGCAGCGGGCTGCGTTTTTCTTT	180				
DB	121	AACACATTTGACGGTTGGAGATAGCAGAGAAAGCAGCGGGCTGCGTTTTTCTTT	180				
QY	181	TTTGTATCCGTTGCGAGCGCAACGAAACGACAGTTGGCAATTCAGACAAAC	240				
DB	181	TTTGTATCCGTTGCGAGCGCAACGAAACGACAGTTGGCAATTCAGACAAAC	240				
QY	241	ACACATCTAAGCCGACCGCAAGCAGCACACACACACACTGCGACACTCGAAGAAA	300				
DB	241	ACACATCTAAGCCGACCGCAAGCAGCACACACACACACTGCGACACTCGAAGAAA	300				
QY	301	AAAAAACAGACGCTGTGCGCGACCTCGAACAGCTTGGTTTCGATTTAGTTGTCAATGC	360				
DB	301	AAAAAACAGACGCTGTGCGCGACCTCGAACAGCTTGGTTTCGATTTAGTTGTCAATGC	360				
QY	361	CTTGAATTCGGTTGCGGGCTTAGTTTCCAAAGTTTATCGTCGTCAGAAACGAAA	420				
DB	361	CTTGAATTCGGTTGCGGGCTTAGTTTCCAAAGTTTATCGTCGTCAGAAACGAAA	420				
QY	421	TAAAAATTTTTCGACCTTAAAAAATCTGACTAAAAATTTGTTTGTATGTTATTTAT	480				
DB	421	TAAAAATTTTTCGACCTTAAAAAATCTGACTAAAAATTTGTTTGTATGTTATTTAT	480				
QY	481	TAGGCACATTTTTCACACACACACACACACTACATCTACAGCTTAAACGGAACCTCTCT	540				
DB	481	TAGGCACATTTTTCACACACACACACACACTACATCTACAGCTTAAACGGAACCTCTCT	540				
QY	541	GCAACAGTGGAGTTGCTGCTCCATCAAGCAGTACTCGAGTTTAAACGAGGATTAACCGG	600				
DB	541	GCAACAGTGGAGTTGCTGCTCCATCAAGCAGTACTCGAGTTTAAACGAGGATTAACCGG	600				
QY	601	GAGAAAGAGAAAGATCGGTGGAGATAGAGATATACAGTGGAGTCAAGAGAGAGGA	660				
DB	601	GAGAAAGAGAAAGATCGGTGGAGATAGAGATATACAGTGGAGTCAAGAGAGAGGA	660				
QY	661	TCATGGACATGATTACGGTGGGGCAGAGCGTCAAGATCAAGCGGACGGATGGCGGTCC	720				
DB	661	TCATGGACATGATTACGGTGGGGCAGAGCGTCAAGATCAAGCGGACGGATGGCGGTCC	720				
QY	721	ACATGGCGGTGGTGGCGGTGATCAACAGTGGGGCAAGTGATCAGTCCAAATGGGTACG	780				
DB	721	ACATGGCGGTGGTGGCGGTGATCAACAGTGGGGCAAGTGATCAGTCCAAATGGGTACG	780				
QY	781	AGCGCGGAAACGAGGAGGAGGAGTGAAGTGGAGCCCATCTACGCTCAATCCGG	840				
DB	781	AGCGCGGAAACGAGGAGGAGGAGTGAAGTGGAGCCCATCTACGCTCAATCCGG	840				
QY	841	AGCTAATGCAAGATATCTGCAACAGCAGCGCGCGCGCGAGCCCAAGAAACAGCCACCG	900				
DB	841	AGCTAATGCAAGATATCTGCAACAGCAGCGCGCGCGCGAGCCCAAGAAACAGCCACCG	900				
QY	901	CGCGATGAACTCTCGGTGTAATCCACACAAATCGGCTATCGGTGGCAATCTCACAGCC	960				
DB	901	CGCGATGAACTCTCGGTGTAATCCACACAAATCGGCTATCGGTGGCAATCTCACAGCC	960				
QY	961	GTATGACCATGCGCGGAAACATGCTGAAACAGATCCAGGAAAGCCAGTCCGATTC	1020				
DB	961	GTATGACCATGCGCGGAAACATGCTGAAACAGATCCAGGAAAGCCAGTCCGATTC	1020				
QY	1021	CGATTTGTCAGCAATAGCGTGAATACAAACAGCAACTCCAACTACGCGCGCGGAG	1080				

DB	1021	CGATTTGTCAGCAGCAATAGCGTGAATACAAACAGCAACTCCCACTACGCGCGCGGAG	1080				
QY	1081	GTGTGGGACCAACAGCTGCGACGACCACTGATTAAGCGTCCAGGTACTCGCAAGCTG	1140				
DB	1081	GTGTGGGACCAACAGCTGCGACGACCACTGATTAAGCGTCCAGGTACTCGCAAGCTG	1140				
QY	1141	CTACGGGCGACGACGACAGCAAGAGTCCGCTCGGCGGTGCTTAATAACATTTGCCAATC	1200				
DB	1141	CTACGGGCGACGACGACAGCAAGAGTCCGCTCGGCGGTGCTTAATAACATTTGCCAATC	1200				
QY	1201	CCAGCGGCGACGACGCTGCTGCTCGGCGGACAAAGAGTCCCACTCGGCGCCCAACCC	1260				
DB	1201	CCAGCGGCGACGACGCTGCTGCTCGGCGGACAAAGAGTCCCACTCGGCGCCCAACCC	1260				
QY	1261	AGGAGCTGGCGGCGCTAGTACCCGCGGATCGCACGCAATGAAAGAGTGGAGGACTGA	1320				
DB	1261	AGGAGCTGGCGGCGCTAGTACCCGCGGATCGCACGCAATGAAAGAGTGGAGGACTGA	1320				
QY	1321	AGGAGATTCGCGAGAGCGCGCCGACAGGCGCGAGATGAAGAGGAGAGGTGGCGC	1380				
DB	1321	AGGAGATTCGCGAGAGCGCGCCGACAGGCGCGAGATGAAGAGGAGAGGTGGCGC	1380				
QY	1381	TGATGAACCGAGATCCGCGCAATCCAACTCGGAGACGCGCCAAATGATACGCAATATC	1440				
DB	1381	TGATGAACCGAGATCCGCGCAATCCAACTCGGAGACGCGCCAAATGATACGCAATATC	1440				
QY	1441	AGAGCACTCGGAATTTGTCGCGCTGCTGATGCGCAGGCGCTGATGACCATCAGATCA	1500				
DB	1441	AGAGCACTCGGAATTTGTCGCGCTGCTGATGCGCAGGCGCTGATGACCATCAGATCA	1500				
QY	1501	CAGTGTGCTGCGCAAGCGTCCCATTAGCCGAGGAGTCAATCGCAAGGAGATCGATG	1560				
DB	1501	CAGTGTGCTGCGCAAGCGTCCCATTAGCCGAGGAGTCAATCGCAAGGAGATCGATG	1560				
QY	1561	TCATTTGCTGCGCGCAAGGACATGCTCATGCTGCAAGGAGTCAATCGCAAGGAGATCGATG	1620				
DB	1561	TCATTTGCTGCGCGCAAGGACATGCTCATGCTGCAAGGAGTCAATCGCAAGGAGATCGATG	1620				
QY	1621	TCACCAAGTTCTCGAGAACCAAGTTTTCGCTTCCGACTAGCCCTTCAACACGACGTCG	1680				
DB	1621	TCACCAAGTTCTCGAGAACCAAGTTTTCGCTTCCGACTAGCCCTTCAACACGACGTCG	1680				
QY	1681	ACAATGCGATGATACAAATACACAGCAAGCGTGGTGAACCAATTTTCAGAGGCG	1740				
DB	1681	ACAATGCGATGATACAAATACACAGCAAGCGTGGTGAACCAATTTTCAGAGGCG	1740				
QY	1741	GAATGCGGACGTGCTTCGCTTACGCGCAGAGGATCGGGCAAAACGACACATTTGGGCG	1800				
DB	1741	GAATGCGGACGTGCTTCGCTTACGCGCAGAGGATCGGGCAAAACGACACATTTGGGCG	1800				
QY	1801	GTGATTTTAAATGGAAGGTGAGACTGCAAGAAACGCGGATCTAGCCATGCGCGCCAGG	1860				
DB	1801	GTGATTTTAAATGGAAGGTGAGACTGCAAGAAACGCGGATCTAGCCATGCGCGCCAGG	1860				
QY	1861	ATGCTTTTGTGACCTTGAATATGCGCGTTACCGCGCCATGAACTAGTCTCGGCCA	1920				
DB	1861	ATGCTTTTGTGACCTTGAATATGCGCGTTACCGCGCCATGAACTAGTCTCGGCCA	1920				
QY	1921	GTTCCTTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTCTGTCGCAAGCAGAACTGC	1980				
DB	1921	GTTCCTTTTGAGATTTACAGTGGCAAGGTCTTTCGATCTTCTGTCGCAAGCAGAACTGC	1980				
QY	1981	CGCTCTCGGAGATGTTAAACAGCAAGTGCAGGTGGTGGGACTCACCGAGAGGTGGTCG	2040				
DB	1981	CGCTCTCGGAGATGTTAAACAGCAAGTGCAGGTGGTGGGACTCACCGAGAGGTGGTCG	2040				
QY	2041	ATGCGCTCGAGGAGTACTGAAGCTCATCCAGCAGCGCAATGCTGCCGCAATCCGCGC	2100				
DB	2041	ATGCGCTCGAGGAGTACTGAAGCTCATCCAGCAGCGCAATGCTGCCGCAATCCGCGC	2100				
QY	2101	AGAGCTCGGCAACTCCAAATTCGTCGCTTCGACGCGCTTTTCCAGATTGTCGCGGC	2160				





```
Db 1959 ATCTACGATAACCGAACTCTCTCTGCAAGCAGTGGAACTGTGTCTCAATCAAGCAGTAC 2018
Qy 576 TCGGAGTTAAACCCAGGATAAGCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATA 635
Db 2019 TCGGAGTTAAACCCAGGATAAGCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATA 2078
Qy 636 TACAGTGGAGTCAAGAGAGAGAGATCATGGACATGATTACGCTGGGCGAGAGCTCAAG 695
Db 2079 TACAGTGGAGTCAAGAGAGAGAGATCATGGACATGATTACGCTGGGCGAGAGCTCAAG 2138
Qy 696 ATCAAGCGGACGGATGGCGGCTCCACATGATGGCCGTGGTGGCGGTGATCAACAGTCGGGC 755
Db 2139 ATCAAGCGGACGGATGGCGGCTCCACATGATGGCCGTGGTGGCGGTGATCAACAGTCGGGC 2198
Qy 756 AAGTGCATCACAGTCCGATGTTACGAGCGCGGGAACGAGAGGCAAGAGGTAGAACTG 815
Db 2199 AAGTGCATCACAGTCCGATGTTACGAGCGCGGGAACGAGAGGCAAGAGGTAGAACTG 2258
Qy 816 GACGCCATCTACGCTCAATCCGAGCTAATGCAAGATCTGTGCAACAGCAGCCCGCC 875
Db 2259 GACGCCATCTACGCTCAATCCGAGCTAATGCAAGATCTGTGCAACAGCAGCCCGCC 2318
Qy 876 CCGGAGCCCAAGAAACAGCCACCGCGCGGATGAACCTCTCGCGTAATCCCAACAATCG 935
Db 2319 CCGGAGCCCAAGAAACAGCCACCGCGCGGATGAACCTCTCGCGTAATCCCAACAATCG 2378
Qy 936 GCTATCGGTGGCAATCTCACCAGCGGTATGACCATGGCGGGAACATGCTGCAACAGATC 995
Db 2379 GCTATCGGTGGCAATCTCACCAGCGGTATGACCATGGCGGGAACATGCTGCAACAGATC 2438
Qy 996 CAGGAAAGCCAGTCCGATTCCTCAATCCGATTTGTGACAGCAATAGCGTGAATACAAACAGC 1055
Db 2439 CAGGAAAGCCAGTCCGATTCCTCAATCCGATTTGTGACAGCAATAGCGTGAATACAAACAGC 2498
Qy 1056 AACTCCAACTACGCGCGGAGGTGGTGGCACCAACAGTGCAGCACCACTGGATTA 1115
Db 2499 AACTCCAACTACGCGCGGAGGTGGTGGCACCAACAGTGCAGCACCACTGGATTA 2558
Qy 1116 CAGCGTCCAGCGTACTCCGAGCTGCTACCGGCGAGCAGCAGAGATGCTCGCG 1175
Db 2559 CAGCGTCCAGCGTACTCCGAGCTGCTACCGGCGAGCAGCAGAGATGCTCGCGCG 2618
Qy 1176 GTGCTTAATAACACTTCCCAATCCAGCGCGGAGCTGCTGTCGGCGGCACAA 1235
Db 2619 GTGCTTAATAACACTTCCCAATCCAGCGCGGAGCTGCTGTCGGCGGCACAA 2678
Qy 1236 GGAGTGGCCACTGCGGCCACAAACCCAGGAGTGGCGCGCTAGTAPACCGCGGATGCGAC 1295
Db 2679 GGAGTGGCCACTGCGGCCACAAACCCAGGAGTGGCGCGCTAGTAPACCGCGGATGCGAC 2738
Qy 1296 GCATTGAAGAGGTGGAGGAGTGAAGAGAGATCGCGAGAGCGAGCGCGCGACAGGCC 1355
Db 2739 GCATTGAAGAGGTGGAGGAGTGAAGAGAGATCGCGAGAGCGAGCGCGCGCGACAGGCC 2798
Qy 1356 GAGATGAAGGAGGAGAGTGGCGCTGATGAACACAGGATCCGGCAATCCAAACTGGGAG 1415
Db 2799 GAGATGAAGGAGGAGAGTGGCGCTGATGAACACAGGATCCGGCAATCCAAACTGGGAG 2858
Qy 1416 ACGGCGCAATGATACGCGCAATATCAGAGCAGCTGGAAATTTGTGCCGCTCTCGATGGC 1475
Db 2859 ACGGCGCAATGATACGCGCAATATCAGAGCAGCTGGAAATTTGTGCCGCTCTCGATGGC 2918
Qy 1476 CAGCGCTGCGATGACCATCAGATCAGATGATGCGTGGCGAGCGCTCCCATAGCGGAG 1535
Db 2919 CAGCGCTGCGATGACCATCAGATCAGATGATGCGTGGCGAGCGCTCCCATAGCGGAG 2978
Qy 1536 GAGTCAATCGCAAGGAGATCGATGCTCATTTTCCGTGCGCGCAAGGACATGCTCATCGTG 1595
Db 2979 GAGTCAATCGCAAGGAGATCGATGCTCATTTTCCGTGCGCGCAAGGACATGCTCATCGTG 3038
Qy 1596 CACGAGCGCGCAGAGGTGCACTCAACAGTTCTTGGAGAAACCAAGTTTCGTGCTTC 1655
```

```
Db 3039 CACGAGCGCGCAGCAGGTCGACCTCACCAAGTTCTCTGGAGAACCAACAAGTTTCGTTC 3098
Qy 1656 GACTACGCTTCAACGACACAGTGGGAGCAATGCCATGTTATACAAATACACAGCCAGCC 1714
Db 3099 GACTACGCTTCAACGACACAGTGGGAGCAATGCCATGTTATACAAATGAGTATCAATCC 3157

RESULT 4
ABL05367
ID ABL05367 standard; cDNA; 2190 BP.
XX
AC ABL05367;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10583.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WI WI; 2001-656860/75.
XX
P-PSDB; ABB61264.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
PS Claim 1; SEQ ID NO 10583; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins (ABBS7737-
XX AB872072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 2190 BP; 586 A; 574 C; 626 G; 404 T; 0 U; 0 Other;

Query Match 18.3%; Score 564.6; DB 4; Length 2190;
Best Local Similarity 68.1%; Pred. No. 3.1e-134;
Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;

Qy 1296 GCATTGAAGAGGTGGAGGAGTGAAGAGAGATCGCGAGAGCGAGCGCGCGACAGGCC 1355
Db 493 GTAGTGAAGGAGGTGAATCGCATGAAGAGCAGAGGAGAGAGCGAAGGGCTCGCAGGCG 552
Qy 1356 GAGATGAAGGAGGAGAGTGGCGCTGATGAACACAGGATCCGGCAATCCAAACTGGGAG 1415
Db 553 GAACAGCTCCAGGAGAGGATGCACTCGCTCGCAATTAATCCGGGGAATCCCAACTGGGAG 612
Qy 1416 ACGGCGCAATGATACCGGCAATATCAGAGCAGCTGGAAATTTGTGCCGCT-----GCTC 1469
Db 613 GTGTGCGTGTGCTGCTGCCCAATACCGCTCCACCTTGATCTTTTCTCCACTTCGATGCTG 672
```

```
QY 1470 GATGGCCAGGCGCTGATGACCATCAGATCAGATGCGTGGCGAAGCGTCCCATTTAGC 1529
D 1471 |||||
D 673 GATCCCAATGAGGCACTGTTAGCAAAATTACGGTGTGTGTGCGAAACACCCATGATG 732
QY 1530 CGCAAGAGGTCAATCGCAAGGAGATCGATGTTCATTTTCGGTGGCGCGCAAGGACATGCTC 1589
D 1531 |||||
D 733 CGCAAGGAGGAGAACTCCAAGAACCTTGGAATCATCATCAGTTCACAGTGGCGGACAGCCTG 792
QY 1590 ATCGTGACAGCGCGGACAGAGGTGCGACCTCAACCAAGTTCCTGGAGAAACCAAGATTT 1649
D 1591 |||||
D 793 ATCGTCCATGATGTGGCGCTCAAGGTGATCTCACCAGTTCCTGGAGCACCAACAATTC 852
QY 1650 CGCTTCGACTACGCTTCAACGACACGTGCGGACCAATGCCATGCTATACAAATACACAGCC 1709
D 1651 |||||
D 853 CGTTTCGACTACAGCTTCGACGAGGAGTGTCTCAATGCGCTGTCTACGATCATCATGCT 912
QY 1710 AAGCCGTGTGTGAAACCATTTTCGAGGCGGGAATGGCAGCTGCTTCGGCTACGGCCAG 1769
D 1711 |||||
D 913 CGTCCGTGTGATCAGAACCATGTTTCGAGGCGGCAATGCCACTTGTTCGCTTACGGACAA 972
QY 1770 ACGGATCGGGCAAAACGCAACCATGGGCGGTGAGTTTAATGGAAGGTGCAAGGACTGC 1829
D 1771 |||||
D 973 ACTGGCAGCGGAAAAACGCAACCATGGGCGGAGAAATTCCTCGGAAAGGTTCAAGGATTC 1032
QY 1830 AAGAACGGCATCTACGCCATGCGGCGGCAAGGATGTCTTTGTGACCTGATATGCCCGT 1889
D 1831 |||||
D 1033 GGTACCGGATCTACGCCATGGCAGCTCGCGATGTCTTCGAGGAGTATCGGCCCGGAG 1092
QY 1890 TACCGGCCCATGAATCTAGTCGTCTCGGCCAGTTCCTTTGAGATTTTACAGTGCAGAGTTC 1949
D 1891 |||||
D 1093 TACCGGCAATGGTGCCAGATTAGTGCAGCTTCTTCGAATCTATGGCACCAGGTG 1152
QY 1950 TTGATC---TTCTGTCCGAAGAGCAAACTCGCGCTCTCGGAGTGTGTAACAGCAA 2006
D 1951 |||||
D 1153 TTGATCTCTCTGCTACCCAAACAGCCCATGCTCGCGGTCTTAGAGGATGCCAGGACAG 1212
QY 2007 GTGCAGGTGGTGGACTCACCGAAGAGTGTGATGCGGTGCGAGAGGTACTGAAGTTC 2066
D 2008 |||||
D 1213 GTCGTGTGTGGTGGCTTAACGGAGATGCGCGGTGACCAAAAGTGGAGATGTCTCTGAGACTG 1272
QY 2067 ATCCAGCAGCGGCAATGCTGCCCAACATCCGGCCAGACGTCCGCCAACTCCCAATTCGTG 2126
D 2068 |||||
D 1273 ATTGACACGGCAGCAAGAGCGCACTTCGGGCCAAACATCGGCGAACCGCAAGTCATCG 1332
QY 2127 CGTTCCGACCGCTTTTCAGATTTGTGCTGCGGCCGACAGGGCTCGACGAAGATCCATGGC 2186
D 2128 |||||
D 1333 CGTTCCGACCGCTCTTTCAAAATAGCACTCCACTTTCCCGATTCTCTGGGGCCACACGGC 1392
QY 2187 AAGTTCTCGTTTCATCGATCTGGGGGCAATGAGCGGGGCGTGGACACTTCTCTGGCCGAT 2246
D 2188 |||||
D 1393 AAGTGTCTCTTTGTGACTTGGCGGGCAATGAACGCGGGGCGGATACGCAATCCGCCGAT 1452
QY 2247 CGGCGACGCGTATGAGGGTGGCGAGATTAAACAAATCGCTGCTGGCCCTCAAGGAGTGC 2306
D 2248 |||||
D 1453 CGTCAAACTCGCATCGAGGGAGCCGAGATCAATAATCTCTGCTGGCCCTCAAGGAGTGC 1512
QY 2307 ATTGCGGCTTGGGCAAAACAGTGGGCCCACTTGCCCTTCCGTGTCTTCCAAATCTCAACCCAG 2366
D 2308 |||||
D 1513 ATTCGAGCCCTCAGCGCGCAGTCGAGTCACCTTCCTTCCTCGTGGCTCCAAAGTTGACCCAA 1572
QY 2367 GTGCTCGCGACTCTGTTCAATTGGCG---AGAAGACGAAGAGTGCATGTATGCGCATGATC 2423
D 2368 |||||
D 1573 GTGCTCGCGACTCTCTTTGTGCGCGCAAGAAGAACAGACCTGCAATGATTCGCATGATA 1632
QY 2424 TCGCGGGAATAGTCTCTCGAGCACACGCTCAACAGCCTGGCTATTCGCGATCGTGTTC 2483
D 2425 |||||
D 1633 TCGCCATCCATGAGCTGCGTGAGAAATACGCTCAACACTCTACGTTACGACAGAGGTT 1692
QY 2484 AAGGAGCTGGTGGTCAAGGATATCG 2508
D 2485 |||||
D 1693 AAGGAGCTCATGCCAAGGAAGACG 1717
```

## RESULT 5

ABL05366  
ID ABL05366 standard; cDNA; 4190 BP.

XX ABL05366;

AC ABL05366;  
XX  
DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 10580.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PS CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

XX P-PSDB; ABB61263.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

PS Claim 1; SEQ ID NO 10580; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 4190 BP; 1187 A; 1021 C; 1025 G; 957 T; 0 U; 0 Other;

Query Match 18.3%; Score 564.6; DB 4; Length 4190;

Best Local Similarity 68.1%; Pred. No. 4.2e-134;

Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;

QY 1296 GCATTGAAGAGGTGGAGCGACTGAAGAGAAATCGGAGAACGCGCCGACAGGCC 1355

Db 1493 GTAGTGAAGAGGTGAATCGCATGAAGGAGCAGAGGAGAGCGAAGGCTGCCAGGCG 1552

QY 1356 GAGATGAAGGAGGAGAGGTGGCGCTGATGACCAAGATCCGGGCNATCCAACTGGGAG 1415

Db 1553 GAACAGCTCCAGGAGAGGATGCACTGCGTCGCAATAATCCGGGGAATCCCACTGGGAG 1612

QY 1416 ACGCGGCAAAATGATACGGCAATATCAGAGCAGCTGGAATTTGTCCGCT-----GCTC 1469

Db 1613 GTGTGCGTGTGCTGGCCNATACCGCTCCACCTTGATCTTCTCCACTTCGATGCCTG 1672

QY 1470 GATGCCAGGCGGTGCGATGACCATCAGATCAGATGTGCGTGCAGAGCGTCCCATTTAGC 1529

Db 1673 GATCCAAATGGAGGCACTGTTTCAGCAAAATACGGTGTGTGTCGGAAGAACGACCCATGAT 1732

QY 1530 CGCAAGGAGGTCAATCGCAGAGGATCGATGTCTTTGGTCCCGCGCAAGACATGCTC 1589

Db 1733 CGCAAGGAGGAGAACTCCAGAACCTGGACATCATCAGATTTCCAGTTCGCGACAGCCTG 1792

1590 ATCGTGCACGACCGCGCAGCAAGGTTCGACCTCACCAGATTCTCTGGAGAACCAAGTTT 1649  
 1793 ATCGTCCATGAGTTGGCTCAAGTGGATCTCACCAGATTCTCTGGAGAACCAAGTTT 1852  
 1650 CGCTTCGATAGCTTCAACACACGACGTCGCAATGCGCATGGTATACAAATACACAGCC 1709  
 1853 CGTTTCGATACACGTCGACGAGGAGTCTCAATGGCTGGTCTACGATCACACTGCT 1912  
 1710 AAGCGTTGGTGAACACATTTTCAGAGGCGGAATGGGAGCTGCTCGCTACGCCAG 1769  
 1913 CGTCCGTTGATCAGAACCATGTTTCAGGCGGCAATGCCACTTGTTCGTTACGGACAA 1972  
 1770 ACGGATCGGCAACACGACACCATGCGCGGTGAGTTTAAATGGAAGGTGCGAGACTGC 1829  
 1973 ACTGCGAGCGGAACCAACACACCATGCGCGGAGATTTCTTCGGAAGGTTTCAGATTGC 2032  
 1830 AAGAACGCGATACGCAATGCGCGCAAGAGATGCTTTGTGACCTGAATATGCCGCT 1889  
 2033 GGTACCGGGATCTACGCCATGCGACTCGCGATGCTTCGAGGAGGTATCGCGCCGAG 2092  
 1890 TACCGCGCATGAATCTAGTCTCGGCCAGTTTCTTTGAGATTTACAGTGGCGAGTTC 1949  
 2093 TACCGCGCAATGGGTGCGCAAGATTACGTGCGACTTCTTCGAAATCTATGGCACCAGGTG 2152  
 1950 TTGCATC---TTCTGTCCGACAGCAAGAACTGCGCGTCTCGGAGATGGTAAACAGCAA 2006  
 2153 TTGCATCTTGTCTACCCCAACAGCCCATGCTCGGGTCTTAGAGATGCGCAGGAG 2212  
 2007 GTGCAAGTGGTGGGACTCACCAGAGAGTGTGCGATGCGGTGCGAGGAGTACTGAAGTTC 2066  
 2213 GTGCGTGGTGGGCTTAAACGAGATGCGGTGACCAAGTGGAGATGCTCTGAGACTG 2272  
 2067 ATCCAGCAGCGCAATGCTGCGCGACATCGCGCAGAGCTGGCCAACTTCAATTCGTG 2126  
 2273 ATTGAGCAGCGGAGCAAGAGCGCACTTCCGGCCAAACATCGCGCAACCGCAAGTCACTG 2332  
 2127 CGTTGCGACGCGCTTTTCCAGATTGTGCTGCGCGCGCAGGGCTCGACGAAGATCCATGGC 2186  
 2333 CGTTCCACGCGCTTTTCAANTAGCATCTCACTTCCCGATTCTGGGGCCACACGGC 2392  
 2187 AAGTTCCTGTTATGATCTCGCGGCAATGAGCGGGCGTGGACACTTCTCGGCCGAT 2246  
 2393 AAGTGTCTCTTTGTGGACTTGGCGGCAATGAACGCGGGCGGATACGCAATCCGCGAT 2452  
 2247 CGGAGACGCTATGAGGCTGCGGAGATTAAACAAATCGCTGCTGGCCCTCAAGAGTGC 2306  
 2453 CGTCAAACTCGATCGAGGGAGCGGAGATCAATAAATCTCTGCTGGCCCTCAAGGAGTGC 2512  
 2307 ATTGCTGCTGGGCAACAGTTCGCGCCCACTTGGCCCTTCCGCTGTCTCCAAACTCACCCAG 2366  
 2513 ATTGAGCCCTCAGCGGCACTCGAGTCACTTCCCTTCCGCTGCTCGAGTTGACCCAA 2572  
 2367 GTGCTGCGGACTCTGTTCAITGGCG---AGAAGAGCAAGAGTGCATGATAGCCATGATC 2423  
 2573 GTGCTGCGGACTCTTGTGCGGCGCAAGAAAGAACAGACCTGATGATGCTCATGATA 2632  
 2424 TCGCGGAGACTTACCTCTGCGACACACGCTCAACAGCTGCGCTGCGGATGCGGTGTC 2483  
 2533 TCGGCATCCATGAGTGTGCTGAGAAATACGCTCAACACTTACGTTACGCGAGACAGGTT 2692  
 2484 AAGGAGCTGTGTGCAAGGATATCG 2508  
 2593 AAGGAGCTATAGCCAGGAGAGCG 2717

RESULT

ABL05833  
 ID ABL05833 standard; cDNA; 1970 BP.  
 AC ABL05833;  
 XX  
 DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 11981.  
 DE  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 OS  
 XX Drosophila melanogaster.  
 PN W0200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 FA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 WIPI; 2001-656860/75.  
 DR P-PSDB; ABB61730.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 XX interactions.  
 PS Claim 1; SEQ ID NO 11981; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC AB872072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1970 BP; 515 A; 521 C; 538 G; 396 T; 0 U; 0 Other;

Query Match 16.4%; Score 505.8; DB 4; Length 1970;  
 Best Local Similarity 64.7%; Pred. No. 3.9e-119;  
 Matches 822; Conservative 0; Mismatches 432; Indels 17; Gaps 4;  
 Qy 1252 CCACAAACCCAGGAGCTGGCGCGCTAGTACCCGCGCATCGCAGCATTTGAAAGAGGTGG 1311  
 Db 399 CCGCCACGCCAGCAGCGCGCGCCAGTTTCGAGAGGA--CGAGGTGTTGCCAGGCTG 456  
 Qy 1312 AGCGACTGAAGGAGATTCGCGAGAGCGAGCGCCCGACAGGCGCGATGAGAGAGAGA 1371  
 Db 457 AAAGAATGCGAAAAGAAACGGGAGCGGAGGAGAGCAAGCCAGGACTCGTCTAGATC 516  
 Qy 1372 AGGTGGCGCTGATGAACACGAGATCCGGGCAATCCAAACTCGGAGACGCGCAATGATAC 1431  
 Db 517 GGGAGCAGGGGAGAGACGAAGATCCGGGAAATCCCAACTGGAGTAGCCAGATGATAC 576  
 Qy 1432 GCGAATATCAGAGCAGCTCGGAATTTGTGCGCTGCTCGATGG-----CCAGGCCG 1482  
 Db 577 GACTGCAACCGAGCAAAATGGAGATCAGCGGTGAGAGTGGTACTACGAACGAACGAA 636  
 Qy 1483 TCGATGACCATCAGATCAGATGTGCGTGCAGAGCGTCCATTTAGCGCGCAGAGAGTCA 1542  
 Db 637 TCAATTTGCCACCAAAATTTATGTTTGTGTGAGGAGAGACCACCTGAGCGCAAGAGTGG 696  
 Qy 1543 ATCGAAGGAGATCGATGTCTATTTTCGTTGCGCGCGCAGGAGCATGCTCATGTGCACGAGC 1602  
 Db 697 CTGACCGGGAACAGGATGTGTCAGCATTCGCTTAGCAACACATTTGTTGTTCCACGAGC 756  
 Qy 1603 CGCGCAGCAAGGTGCGACCTTCCCAAGTTCTCGGAGAACCAAGTTTCGTTTCGATACG 1662

```
Db 757 CCGCAAGCATGTGAACCTGGTCAAGTTCTCTGAAATCATAGCTTCCTGATTACG 816
Qy 1663 CTTTAAACAGACGTGCGCAATATGCGCATGTATACAAATACACAGCAAGCCGTTGGTGA 1722
Db 817 TCTTCAGCAGGAGTCTCCAATGCCACGGTCTACGAAATTCACAGCCGACCTTGATAA 876
Qy 1723 AAACCAATTTTCGAGGCGGAAATGGGCAAGTGTCTGCGCTACGCGCAGCGGATCGGCA 1782
Db 877 AGCACATTTTGTATGGCGGAATGGCCACGTTCTGCGCTACGCGCAAACTGGAAGCGCA 936
Qy 1783 AAACGCACACCATGGCGCGTGAATTTAATGGAAGGTGAGGATCTCAAGAACCGCATCT 1842
Db 937 AGACCTATACGATGGTGTGTCAGTTCCCGGAAGGATCAGAGCTCAATGATGGCATCT 996
Qy 1843 AGCCATGGCGGCAAGGATGCTTTGTGACCTGAATATGCGCGTTTACCGCGCCATGA 1902
Db 997 ATGCAATGGCGCTAAGGACGTGTTCTCCACTCTAAAGACGGTTCCTATACAAAGCTTA 1056
Qy 1903 ATCTAGTCTCTCGGCCAGTTCTTTTGAGATTTACAGTGGCAAGTCTTCGATCTTC--- 1959
Db 1057 ATCTGAAAGTTTACTGCGACGTTCTTCGAGATCTACGGCACCCCGGGTGTTCGATCTCTGA 1116
Qy 1960 TGTCCGACAAAGCAGAACTGCGCTCTCGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019
Db 1117 TGCCTGGCAAGCCACACTGCGTGTCTTGAGATACAGGATAGAACAGGAGTGCAGGTGG 1176
Qy 2020 GACTCACCGAAGGTGTCGATGGCGTCGAGGAGTACTGAAGCTCATCCAGCACGGCA 2079
Db 1177 GCCTCACCCAGAATCCAGTACAGAACACCGCGAAGTTCTGACCTACTCGAGTTGGCA 1236
Qy 2080 ATGCTGCCGACATCCGCGCAGAGTGGCCCACTCCAAATTCGTGGGTTCGCAAGCGG 2139
Db 1237 ATAGTGTCCGAACCTCGGGTCAACCTCTGCCAATTTCCAAGTCTCCCGATCGCATGCTG 1296
Qy 2140 TTTTCCAGATTTGCTGCGCGCGCAGGGCTCGACGAAGATCCATGSCAAGTTCTGTTCA 2199
Db 1297 TGTTCCAAATCGTGTGAGATCCGCGGGCGAGAGTACACGGGAAATTCCTGCTTA 1356
Qy 2200 TCGATCTGGCGGCAATAGCGGGGCTGGACACTTCTCGCCGATCGGCAGACGGTA 2259
Db 1357 TAGATCTGGCGGGAATGAAGAGGAGCGGACACAGCTCGCGGATCGACAGACGGCC 1416
Qy 2260 TGGAGGTGCGGAGATTAACAATCGTCTGGCCCTCAGAGGTGCAATTCGTGGTGG 2319
Db 1417 TGGAGGATCCGAGATCAATAAATCGCTGCTGCTCAAGGAATGCAATTCGCGTCTGG 1476
Qy 2320 GCAAAAGTTCGCGCCACTTTCCTTCCTGCTCCAACTCACCCAGGTGCTGCGCGACT 2379
Db 1477 GCGCGCAGTGGATCATTTTGCATTTCCGTGGCTCCAAGTGAACCAAGTCTGCGGGACT 1536
Qy 2380 CTTTCATTTGGCG---AGAAGAGCAAGACGTGCAATGATAGCCATGATCTCGCGGACTTA 2436
Db 1537 CTTTCATCGGAGTAAGAGGTGAAACCTTGATGATAGCCATGATCTCGCCATGCTTGC 1596
Qy 2437 GCTCTCGGAGCACAGCTCAACAGCTGCGGTATGCGGATGCTGTCAAGGAGTGGTGG 2496
Db 1597 ATTTCGTTGAGCATACCTTTGAACACGCTGCGTTATGCGGATCGGTTGAAGGAACCTAAGT 1656
Qy 2497 TCAAGGATATC 2507
Db 1657 TGGATCGATC 1667
```

## RESULT 7

ABL05832/c  
ID ABL05832 standard; cDNA; 3970 BP.

XX  
AC ABL05832;

XX  
DT 26-MAR-2002 (first entry)

XX  
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 11978.

```
Qy 1252 CCACAACCCAGGAGCTGGCGGCTAGTACCCGCGGATCGCACCATTTGAAAGAGGTGG 1311
Db 2572 CCGCACCGCAGCAGACGGCGCGCCAGTTCGAGGGA--CGAGTGTGTGACAGGCTG 2515
Qy 1312 AGCGACTGAAGAGAAATCGCAGAGGACGCGCCCGACAGGCCGAGATGAAGAGGAGA 1371
Db 2514 AAAGAAATCGGAAAAGAACGGGACGAGGCGGAGAACAGCCAGGACTCGTCTAGATC 2455
Qy 1372 AGTGGCGCTGATGAACGAGGATCGGGCATCCAACTGGGAGACGGCGCAATGATAC 1431
Db 2454 GGGAGCAGGGGAAGAACGAAATCCGGGAAATCCCACTGGGAAAGTAGCCAGAATGATAC 2395
Qy 1432 CGGAATATCAGACACGCTGGAATTTGTGCGCTGCTCGATGG-----CCAGGCG 1482
Db 2394 GACTGCAACGCGAGCAATGGAGAGTCAAGCGGTGAGAGTGTACTAGCAACGACGNA 2335
Qy 1483 TCGATGACCATCAGATCAGTGTGCGTGCGCAAGCGTCCATTAGCCGCAAGAGGTGA 1542
Db 2334 TCAATTCGCAACCAATATGTTGTTGTGAGGAGAGACCACTGAGGCGCAAGGAGCTGG 2275
Qy 1543 ATCGCAAGGAGATCGATGTCAATTTGCGTGGCGCGCAGGACATGCTCATGTCGACGAC 1602
Db 2274 CTGACCGGGAAACAGGATGTGGTCAGCATTCCTTAAGCACACATTTGGTGGTCCACGAC 2215
Qy 1603 CGCGCAGCAAGTTCGACTCAACCAAGTTCTCTGGAGAACCAACAGTTTCGCTTCGACTACG 1662
Db 2214 CCGCAAGCATGTGAACCTGCTCAAGTTCTCTGGAAATCATAGTTCGCTTCGATTACG 2155
Qy 1663 CTTTCAACGACACGTCGCGCAATGCCATGGTATACAAATACACAGCAAGCCGTTGGTGA 1722
```

Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical; gene; ss.  
Drosophila melanogaster.  
WO200171042-A2.  
27-SEP-2001.  
23-MAR-2001; 2001WO-US009231.  
23-MAR-2000; 2000US-0191637P.  
11-JUL-2000; 2000US-00614150.  
(PEKE ) PE CORP NY.  
Venter JC, Adams M, Li PWD, Myers EW;  
WPI; 2001-656860/75.  
P-PSDB; ABB61729.  
New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signaling and cell-cell  
interactions.  
Claim 1; SEQ ID NO 11978; 21pp + Sequence Listing; English.  
The invention relates to an isolated nucleic acid detection reagent  
capable of detecting 1000 or more genes from Drosophila. The invention is  
useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
ABB72072). The sequence data for this patent did not form part of the  
printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
Sequence 3970 BP; 969 A; 935 C; 968 G; 1098 T; 0 U; 0 Other;  
Query Match 16.4%; Score 505.8; DB 4; Length 3970;  
Best Local Similarity 64.7%; Pred. No. 5.6e-119;  
Matches 822; Conservative 0; Mismatches 432; Indels 17; Gaps 4;





```
Db 376 TTGATTTGATGAACAGCTTCGAATGAAGTTGTCTACAGTTTCACAGCAAGGCCACTGG 435
Qy 1720 TGAACACATTTTCGAGGGGGAATGCGGACGTCTTCGCTTACGCGGACGAGCGGATCGG 1779
Db 436 TACAGACAATCTTTGAAGGTGGAAGAACACTTTGTTGCTATATGCGCAGACAGGAAGTG 495
Qy 1780 GCAAAACGACACCATCGGCGGTGAGTTTAATGGAAGGTGCAAGACTGCAAGAACGGCA 1839
Db 496 GCAAGACACATATCTGCGGCGGAGACCTCTCTGGAAGGCCAGAAATGCATCAAAAGGGA 555
Qy 1840 TCTACGCCATGCGCGGCCCAAGGATGCTTTGTTGACCTGACCTGATATGCGCGTTACCGGCCA 1899
Db 556 TCTATGCGCATGCGCTCCGCGGAGCTTTCTCTGGAAGATCAACCTGCTACCGGAAGT 615
Qy 1900 TGAATCTAGTCTGCTCGGCCAGTTTCTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC 1959
Db 616 TGGCGCTGGAAGTCTATGTGACATCTCTCGAGATCTACATGGGAAGCTGTTTGACTGTC 675
Qy 1960 TGTCCGACAGCAGAAATCGCGGCTGAGAGATGTTAAACAGCAAGTGCAGGTGGTG 2019
Db 676 TCAACAGAGGCAAGCTCGCGGCTGCGAGGATGCGCAAGCAACAGGTGCAAGTGGTG 735
Qy 2020 GACTCACCGGAGAGGTGCTGATGGGTGCGAGGCTGCTGAGCTCATCCACGCGCA 2079
Db 736 GCGTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGCA 795
Qy 2080 ATGCTGCCGGAACATCGCGGACAGCTCGCGCCAACTCCAAATTCGTCGCGTTCCGACCGG 2139
Db 796 GCGCTGCAGAACCTCTGGGACAGCATTTGCCAACTCCAAATTCCTCCGCTCCACGCGT 855
Qy 2140 TTTTCAGATGCTGCGGCGCGAGGCTGCGAGAGATPCATGCGCAAGTTCTCTTTCA 2199
Db 856 GCTTCCAAATTAATCTTCGAGCTAAAG-----GAGAAATGCATGCGCAAGTTCTCTTTGG 909
Qy 2200 TCGATCTGCGGCAATGAGCGGCGTGGACACTTCTCGCGGATCGCGACACGCTA 2259
Db 910 TAGATCTGGCAGGAATGAGCGGCGGACACTTCCAGTGTGCTGACCGGACAGCCGCA 969
Qy 2260 TGGAGGTGCGGAGATTAACAAATCGCTGCTGCGCTCAAGGAGTGCATTCGTGCGTTGG 2319
Db 970 TGGAGGCGGAGAAATCAACAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGCGCTGG 1029
Qy 2320 GCAACAGTGGCGCCACTTGCCTTCGCTGTCTCCAAATCTCAACCCAGGTGCTGCGGACT 2379
Db 1030 GACAGAACAGGCTCACACCCGCTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1089
Qy 2380 CATTTCATTCGCGAGAGAGCAGACGTGATGATGATGATGATGATGATGATGATGATGAT 2439
Db 1090 CTTTCATTCGCGGAGAACTCTAGGACTTGCATGATTCGCCAGATCTCACCAGGCATAGCT 1149
Qy 2440 CTTGCGAGCACAGCTCAACACGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492
Db 1150 CTTGTGAATATACTTTAAACACCTCGAGATGTCAGACAGGCTCAAGGAGCTG 1202
```

## RESULT 9

ADE10082

ID ADE10082 standard; DNA; 1292 BP.

XX AC ADE10082;

XX AC ADE10082;

DT 29-JAN-2004 (first entry)

DE DNA encoding human MCAK enzyme fragment #1.

XX Human, motor protein; mitotic centromere-associated kinesin; MCAK;  
KW microtubule-stimulated ATPase activity; cellular proliferation disorder;  
KW cancer; autoimmune disease; arthritis; inflammatory bowel disease;  
KW keloid; psoriasis; tumour; cytostatic; immunosuppressive; antiarthritic;  
KW antiinflammatory; gastrointestinal; vulnary; antipsoriatic; ds.  
XX Homo sapiens.

XX

FH

FT

FT

FT

FT

FT

XX

PN

XX

XX

PD

XX

PF

XX

PR

XX

PA

XX

PI

XX

DR

DR

XX

PT

PT

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Location/Qualifiers

1. .1292

/tag= a

/transl\_except= (pos:4..5,aa:Arg)

/note= "this codon has an apparent 1 nucleotide deletion which alters the reading frame"

US6638754-B1.

28-OCT-2003.

29-MAR-2002; 2002US-00112432.

28-NOV-2000; 2000US-00724215.

(CYTO-) CYTOKINETICS INC.

Beraud C, Sakowicz R;

WPI; 2003-842789/78.

P-PSDB; ADE10083.

New nucleic acid encoding motor proteins, useful in identifying compounds for the treatment of cellular proliferation disorders, such as cancer, autoimmune disease, inflammatory bowel disease, arthritis, keloids and psoriasis.

Claim 2; Fig 1; 44pp; English.

The present invention relates to the isolation of polynucleotide sequences encoding fragments of human motor protein, mitotic centromere-associated kinesin (MCAK). The MCAK enzyme fragments have microtubule-stimulated ATPase activity. The invention also discloses methods for the use of the motor protein fragments. The methods and compositions are useful for high throughput screening systems for identifying compounds useful in the treatment of cellular proliferation disorders, such as cancer, autoimmune disease, arthritis, inflammatory bowel disease, keloids, psoriasis and tumours. The present sequence encodes a human MCAK enzyme fragment.

Sequence 1292 BP; 374 A; 287 C; 349 G; 282 T; 0 U; 0 Other;

Query Match

Best Local Similarity 14.8%; Score 454.6; DB 10; Length 1292;

Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

Qy 1300 TGAAGAGGTGAGGAGCTGAAGGAGATCGGAGAGCGACGCGCCACAGGCGGAGA 1359

Db 22 TGAAGGAAGTGGAAAAAATGAAGAACAGCGAGAGAGAAGAGGCCAGACTCTGAAA 81

Qy 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACAGGATCCGGGCAATCCAAACTGGGAGACGG 1419

Db 82 TGAGATGAGAGAG-----CTCAGGAGTATGACAGTAGTATTTCCAACTGGGAATTG 135

Qy 1420 CGCAATGATACCGAATATCAGAGCAGCTGGAATTTGTGCGGCTGCTCGATGGCCAGG 1479

Db 136 CCGGAATGATTAAGAAATTTTCGGGCTACTTTGGAATGTCTACCTACTATGACTGATC 195

Qy 1480 CCGTGTGATGACCATCAGATCAGTGTGGTGGCAGAGGTCCCATTAGCCGCAAGAGG 1539

Db 196 CTATCGAAGAGCAGAAATATGTCTGTGTAGGAAACGCCCACTGAATAAGCAAGAAT 255

Qy 1540 TCAATCGCAAGGAGATCGATGTCATTTTCGGTCCGCGCAAGGACATGCTCATCGTCACG 1599

Db 256 TGGCCAGGAAGAAATTAATGATGATTTTCAATTCCTAGCAAGTGTCTCTCTTGGTACATG 315

Qy 1600 AGCCGCGCAGCAGGTGCGACCTCCACCAAGTTCTCTGGAGAACCAAGTTTCGCTTCGACT 1659

Db 316 RAACCAAGTTGAAGTGGACTTTAAACAAGTATCTGGAGAACCAAGCAATTCGCTTTGACT 375

Qy 1660 AGCCCTTCAACGACACGTCGCGCAATGCCATGGTATACAAATACACAGCCAGCCGTTGG 1719

376	TTGCATTGTAGTAACACAGCTTCGAATGAAGTTGTCTACAGGTTCCACAGCAAGGCCACTGG	435
1720	TGAAAAACCAATTTTCAGAGGCGGAATTGGCGACGTCGCTTCGCCTACGGCCACAGCGGAATCGG	1779
436	TACAGACAATCTTTGAAGGTGGAAAAAGCAACTGTTGTTTGCATATGGCCACAGACAGGAATG	495
1780	GCAAAACGCACACCATGGGCGGTGAGTTTAATGAAAGGTCGAGGACTGCAAGAACGGCA	1839
496	GCAAGACACATACTATGCGGCGGAGACCTCTCTGGGAAAGCCCAAAATGTCATCCAAAGGGA	555
1840	TCTACGCCATCGGCGCAAGGATGTCCTTTGTGACCCCTGAATATATGCCGGTTACCGCGCCA	1899
556	TCTATGCCATGCGCTCCCGGAGCGTCTTCTCTCGAAGATTCACCTGCTACCGGAAGT	615
1900	TGAATCTAGTCGTCTCGGCCAGTTCTTTTGAGATTTACAGTGGCAAGGTCCTTCGATCTTC	1959
616	TGGGCTTGGAAATCTATGTGACATCTTCGAGATCTACAATGGGAAGCTGTTTGACCTGC	675
1960	TGTCGACACAGCAGAAACTGCGGCTCTGAGGATGTTAAACAGCAGTGCAGGTGGTGG	2019
676	TCAACAAGAGGCCCAAGTCGCGCTGCTGGAGGATGGCAAGCAACAGTGCAAAGTGGTGG	735
2020	GACTCACCGAGAAGGTGTCATGGCGCTCGAGGAGGTACTGAAAGCTCATCCAGCAGCGCA	2079
736	GGCTGCGAGGACATCTGGTTAACTCTGCTGATGATGTCATCAGATGATCGATGGGCA	795
2080	ATGTGCGCCGAACATCCGGCGACAGCTCGGCCAACTCCAATTCGTGCGCTTCGCACGCGG	2139
796	GCGGCTGCAGAACCTCTCTGGGAGACATTTGCCAACTCCAATTCCTCCGCTCCACGCGT	855
2140	TTTTTCCAGATTGCTCGGCGCGCAGGCTTCGACGAGATGCCATGGCAAGTTCTCTGTTCA	2199
856	GCITTCACAAATATTTCTTCGAGCTAAAGG-----GAGAAATGCATGGCAAGTTCTCTTTGG	909
2200	TCGATCTGGGGGCAATAGCGGGCGGTGCACACTCTCTCGGCCGATTCGCAGCAGCGGTA	2259
910	TAGATCTGGCAGGAATGAGCGAGGCGGAGACATTCAGTGTCTGACCGGACAGCCGCA	969
2260	TGGAGGTGCGAGATTAAACAATTCGCTGTGGCCCTCAAGGAGTGCAATTCGTGCGTTGG	2319
970	TGGAGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGGAGTGTCATCAGGGCCCTGG	1029
2320	GCAACAGTGGGCCCACTTCGCCCTTCGCTCTCCAAACTCACCCAGGTGCTGCGGACT	2379
1030	GACAGAACAGGCTTCAACCCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTCAGGGACT	1089
2380	CGTTCAATGGCGAAGACGATGCATGATAGCCATGATCTCCCGGGACTTTAGCT	2439
1090	CTTTCATTTGGGAGAACTCTTAGGACTTGCATGATTGCCACGATCTCACCGGCAATAGCT	1149
2440	CTTCGAGCAGCAGCTCAACAGCTCGGCTATGCGGATCGTGTCAAGGAGCTG	2492
1150	CTGTGTAATATATCTTTAAACACCCCTGAGATATGCAGACGGTCAAGGAGCTG	1202

RESULT	ID	DB	Sequence
ADOS2609	196	Db	CTATCGAAGAGCACAGAATATGTGTCTGTGTTAGGAAAACGCCCACTGAATAAGCAAGAAT 255
ADP53609 standard; DNA; 1292 BP.	1540	Qy	TCAATCGCAAGGAGATCGGATGTCAATTCGGTCCCGCAAGGACATGTCTATCGTGCACG 1599
ADP53609;	256	Db	TGGCCAAAGAAAGAAATTTGATGTGATTTCCATTTCTTAGCAAGTGTCTCTCTTGGTACATG 315
12-AUG-2004 (first entry)	1600	Qy	AGCGGCGAGCAGGTCGACCTCACCAAGTTCTCGAGAACCAACAAGTTTCGCTTCGACT 1659
Human mitotic centromere-associated kinesin (MCAK) DNA #1.	316	Db	AACCCCAAGTTGAAAGTGGACTTTACAAAGATATCTCGAGAACCAAGCAATTCGCTTTGACT 375
Mitotic centromere-associated kinesin; MCAK;	1660	Qy	AGCCTTTCAACGACACAGTCGCGACAATGCCATGGTATACAAATACACAGCCAGCGGTTGG 1719
cellular proliferative disease; cancer; hyperplasia; restenosis;	376	Db	TTGCAATTTGATGAAACACAGCTTCGAAATGAAGTTGTCTACAGGTTTCACAGCAAGGCCACTGG 435
cardiac hypertrophy; immune disorder; inflammation; cytostatic;	1720	Qy	TGAAAACCACTTTTCGAGGGCGGAATGGGGAGCTGCTTCGCCATACGCCCCAGACGGGATCGG 1779
vasodilator; therapy; human; gene; ds.	436	Db	TACAGACAATCTTTTGAAGGTGGAAGAAACAACTTTGTTTTCATATGCCCAGACAGGAAGTG 495
Homio sapiens.			
Key			Location/Qualifiers

```
QY 1780 GCAAAAGCGCACACCATGGCGGTGAGTTTAATGGAAGGTGCAAGGACTGCAAGAACGGCA 1839
Db 496 GCAAGACACATATATGGGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCAATCCAAAGGGA 555
QY 1840 TCTACGCCATGGCGGCAAGGATGCTTTGTGACCCCTGGAATATGCCGCGTTACCGCGCCA 1899
Db 556 TCTATGCCATGGCTCCCGGAGCTTCTCTCTGAAGAAATCAACCCTGCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTGCTCGGCGAGTTTCTTTGAGATTTACAGTGGCAAGGTCTTGATCTTTC 1959
Db 616 TGGCGCTTGAAGTCTATGTGACATCTTTCAGATCTACAATGGGAAGCTGTTTGACCTGC 675
QY 1960 TGTCCGACAGCAGAACTGCGCGTCTCGAGAGTGTAAACAGCAAGTGCAGGTGCTGG 2019
Db 676 TCACCAAGAGGCAAGCTGCGCGTCTCGAGAGTGCAGCAAGGTGCAAGTGTGTGG 735
QY 2020 GACTCACCGAGAAGGTGGTTCGATGGCGTTCGAGAGGTACTGAAGCTCATCCAGCAGGCA 2079
Db 736 GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATGATGATGATGATGATGATG 795
QY 2080 ATGCTGCCGAACATCCGCGCAGACGTGCGGCCAACTTCCAATTTGTCGCGTTTCGACGCG 2139
Db 796 GCGCTTGCAGAACCTCTGCGCAGACATTTGCCAACTTCCAATTTCTCCCGCTCCAGCGCT 855
QY 2140 TTTTCCAGATTGTCTGCGCGCGCAGGCTCGACGAGATCCATGGCAAGTCTCGTTCA 2199
Db 856 GTTTCCAAATTTATCTTCGAGCTAAAG-----GAGAAATGATGCAAGTTCTCTTTGG 909
QY 2200 TCGATCTGCGCGCAATGAGCGGGGTGAGACACTTCTCGCGCATCGGCGAGCGGTA 2259
Db 910 TAGATCTGCGAGGAATGAGCGGGCGGACACTTCCAGTGTGACCGCGAGACCGGCA 969
QY 2260 TGGAGGGTCCGAGATTAAACAATCGCTCTGCGCCCTCAAGAGTGCATTTCTGCGGTGG 2319
Db 970 TGGAGGGCGAGAAATCAACAAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGCGCTGG 1029
QY 2320 GCAACAGTGGCGCCACTTGGCCCTCCGTTCTCCAACTCACCGAGTGTGCGGACT 2379
Db 1030 GACGAAACAGGCTCACACCCGTTCCGTGAGAGCAAGCTGACAGAGTGTGAGGGACT 1089
QY 2380 CTTTCAATGGCGAGAGCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2439
Db 1090 CTTTCAATGGCGAGAACTCTAGACTTGTGATGATGATGATGATGATGATGATGATGATGAT 1149
QY 2440 CTGCGAGCACAGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492
Db 1150 CTTGTGAATATACTTTAAACACCTGAGATATGACAGAGGCTCAAGGAGCTG 1202

RESULT 11
AAD24083
ID AAD24083 standard; DNA; 1421 BP.
XX
AC AAD24083;
XX
DT 09-APR-2002 (first entry)
XX
DE Human mitotic centromere-associated kinesin protein fragment #4 DNA.
XX
KW Human; MCAK; mitotic centromere-associated kinesin; motor domain;
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
KW cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; ds.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH 1..1421
FT /*tag= a
FT /*product= "MCAK enzyme fragment"
FT /*transl_except= (pos:4..5, aa:Arg)
FT misc_feature 4
FT /*tag= b
```

/note= "Base G at this location is missing in the sequence shown as SEQ ID NO: 7 in sequence listing of the specification"

US6331424-B1.

18-DEC-2001.

15-JUN-2000; 2000US-00594669.

20-APR-1999; 99US-00295612.

18-MAY-1999; 99US-00314464.

(CYTO-) CYTOKINETICS INC.

Beraud C, Sakowicz R;

WPI; 2002-089075/12.

P-PSDB; AAE14502.

New human MCAK (mitotic centromere-associated kinesin) protein useful in identifying agents for use in the treatment of cellular proliferation disorders.

Disclosure; Fig 7; 44pp; English.

The invention relates to human MCAK (mitotic centromere-associated kinesin) protein, and its fragments that comprise a motor domain and directly or indirectly produce ADP or phosphate. The MCAK enzyme and its fragments are used in methods to identify compounds that modulate their activity. Modulators of MCAK are useful as therapeutic agents for treating cellular proliferation disorders such as cancer, hyperplasia, restenosis, cardiac hypertrophy, immune disorders, inflammation, autoimmune disease, arthritis, graft rejection, inflammatory bowel disease. The present sequence is a human MCAK fragment encoding DNA

Sequence 1421 BP; 417 A; 313 C; 387 G; 304 T; 0 U; 0 Other;

Query Match 14.8%; Score 454.6; DB 6; Length 1421;

Best Local Similarity 62.6%; Pred. No. 5e-106;

Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGCGACTGAAGAGAAATCGGAGAGCGACGCGCCGACAGCGGCGAGA 1359

Db 22 TGAAGAGGTGGAAGAAATGAAGAAACAGCGAGAGAGAGAGAGCCCAAGAACTCTGAAA 81

QY 1360 TGAAGAGAGAGAGTGGCGCTGATGAACAGGATCGGGCAATCAAACTGGGAGACGG 1419

Db 82 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTCTTCCAACTGGGAATTTG 135

QY 1420 CGCAATGATACGCGAATATCAGAGCACGCTGGAAATTTGTGCGCTGCTCGATGGCCAGG 1479

Db 136 CCGAATGATTAAGAAATTTGGGCTACTTTGGAATGTCATCCACTTACTATGACTGATC 195

QY 1480 CCGTGCATGACCATCAGATCAGTGTGCGTGGCAAGCGTCCCAATTAGCCGCAAGGAGG 1539

Db 196 CTATCGAAGAGACACAGAAATATGTCTGTGTAGGAAACGCCCACTGAATAAGCAAGAT 255

QY 1540 TCAATCGCAAGAGATCGATGTCAATTCGGTGGCGGCAAGAGATGCTCATCTGTCAGC 1599

Db 256 TGGCAAGAAAGAAATTTGATGATTTCCATTCCTAGCAAGTGTCTCTCTTTGGTACATG 315

QY 1600 AGCGCGCAGCAAGGTCCACCTCAACAAAGTTCTCGAGAACCAACAAGTTTCGCTTCGACT 1659

Db 316 AACCAAGTTGAAAGTGGACTTAACAAAGTATCTGGAGAACCAAGACTTCTGCTTTGACT 375

QY 1660 AGCGCTTCAACGACAGCTGCGCAATGCTGGTATACAAATPACACAGCAAGCGTTGG 1719

Db 376 TTGCATTTGATGAAACAGCTTTCGAATGAAGTTGCTACAGGTTACACAGAGGCCACTGG 435

QY 1720 TGAAGAACCATTTTCAGGCGCGAATGGGAGAGTGTCTTCGCTACGCGCAGACGGATCGG 1779

Db 436 TACAGCAATCTTTGAAGGTGGAAAGCAACTTGTGTTGATATGCGCCAGACAGGAAGTG 495

QY 1780 GCACAAACGACACCATGGCGGTGAGTTTAAATGGAAAGTGTCAGGACTGCAAGACGGCA 1839  
 Db 496 GCAAGACACATATATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGGA 555  
 QY 1840 TCTACGCAATGGCGGCCAAGAGATGCTTTGTGACCCCTGAATATGCGCGTTACCGCGCA 1899  
 Db 556 TCTATGCCATGGCTCCCGGAGCGTCTTCTCTCTGAAGAATCAACCTGCTACCGGAAGT 615  
 QY 1900 TGATCTAGTCTGCGCCAGTCTTTTGTGAGATTTACAGTGGCAAGCTTTCGATCTTC 1959  
 Db 616 TGGGCTCGAAGTCTATGTGACATCTTCGAGATCTCAATGGGAAGCTTTTGACCTGCG 675  
 QY 1960 TGTCCGACAGCAGAAATCGCGCTCTGGAGATGTTAAACAGCAAGTGCAGGTGGTGG 2019  
 Db 676 TCAACAGAGGCCAAGCTGCGCTGCTGGAGATGGCAGCAAGGTGCAAGTGGTGG 735  
 QY 2020 GACTCACCAGAGAGTGTGATGCGGCTCGAGGAGTACTGAAGTCTCATCAGCACGGCA 2079  
 Db 736 GCGTCGAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATGACATGGGCA 795  
 QY 2080 ATGTCGCCGACATCCGCGCCAGAGTGGCCCACTCCAATTCGTCGGTTCGACGCGG 2139  
 Db 796 GCGCTCGAGAACCTCTGGGAGACATTTGGCAACTCCAAATTCCTCCCGCTCCACGCGT 855  
 QY 2140 TTTTCCAGATGCTGCGCGCGCAGGCTCGACGAAGTCCATGGCAAGTTCGTTCA 2199  
 Db 856 GCTTCCAAATATCTTCGAGTAAAGG-----GAGAAATGATGGAAATTCCTTTGG 909  
 QY 2200 TCGATCTCGCGGCAATGAGCGGGCGTGGACACTTCTCGCGCATCGGAGACGGTA 2259  
 Db 910 TAGATCTGGCAGGAAATGAGCGCGGACACTTCCAGTGTGACCGGAGACCCGCA 969  
 QY 2260 TGGAGGTGCGGAGTTAACTAATCGCTGCTGGCCCTCAAGAGTGCATTCGTGGTGG 2319  
 Db 970 TGGAGGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGCGCCCTG 1029  
 QY 2320 GCAAAACAGTGGCGCCACTTGCCTTCCGTGCTCCAACTCACCAGGTGCTGGCGACT 2379  
 Db 1030 GACAGCAAGAGTCAACCCCGTCTCGTGAAGCAAGTGCACAGGTGCTGAGGACT 1089  
 QY 2380 CGTTCATTTGGGAGAACTCTAGGACTTGATGATGATGATGATGATGATGATGATGAT 2439  
 Db 1090 CTTTCATTTGGGAGAACTCTAGGACTTGATGATGATGATGATGATGATGATGATGAT 1149  
 QY 2440 CTTGGAGCACAGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGAGCTG 2492  
 Db 1150 CTTGTGAATACTTTTAAACACCCCTGAGATATGAGACAGGCTCAAGAGCTG 1202

RESULT 12  
 ADE10088  
 ID AD10088 standard; DNA; 1421 BP.  
 XX AC AD10088;  
 XX AC AD10088;  
 DT 29 JAN-2004 (first entry)  
 XX DE DNA encoding human MCAK enzyme fragment #4.

Human; motor protein; mitotic centromere-associated kinesin; MCAK;  
 microtubule-stimulated ATPase activity; cellular proliferation disorder;  
 cancer; autoimmune disease; arthritis; inflammatory bowel disease;  
 keloid; psoriasis; tumour; cystostatic; immunosuppressive; antiarthritis;  
 antiinflammatory; gastrointestinal; vulnery; antipsoriatic; ds.  
 OS Homo sapiens.  
 XX XX Location/Qualifiers  
 XX XX 1..1421 a  
 XX XX /\*tag=  
 XX XX /transl\_except= (pos:4..5,aa:Arg)  
 XX XX /note= "This codon has an apparent 1 nucleotide deletion

FT XX which alters the reading frame"  
 XX US6638754-B1.  
 XX 28-OCT-2003.  
 PD 29-MAR-2002; 2002US-00112432.  
 XX 28-NOV-2000; 2000US-00724215.  
 PR (CYTO-) CYTOKINETICS INC.  
 XX Beraud C, Sakowicz R;  
 PI WPI; 2003-842789/78.  
 XX P-PSDB; ADE10089.  
 DR  
 XX New nucleic acid encoding motor proteins, useful in identifying compounds  
 PT for the treatment of cellular proliferation disorders, such as cancer,  
 PT autoimmune disease, inflammatory bowel disease, arthritis, keloids and  
 PT psoriasis.  
 XX Claim 2; Fig 7; 44pp; English.  
 PS The present invention relates to the isolation of polynucleotide  
 CC sequences encoding fragments of human motor protein, mitotic centromere-  
 CC associated kinesin (MCAK). The MCAK enzyme fragments have microtubule-  
 CC stimulated ATPase activity. The invention also discloses methods for the  
 CC use of the motor protein fragments. The methods and compositions are  
 CC useful for high throughput screening systems for identifying compounds  
 CC useful in the treatment of cellular proliferation disorders, such as  
 CC cancer, autoimmune disease, arthritis, inflammatory bowel disease,  
 CC keloids, psoriasis and tumours. The present sequence encodes a human MCAK  
 CC enzyme fragment.  
 XX Sequence 1421 BP; 417 A; 313 C; 387 G; 304 T; 0 U; 0 Other;

Query Match 14.88; Score 454.6; DB 10; Length 1421;  
 Best Local Similarity 62.6%; Pred. No. 5e-106;  
 Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
 QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGAAATCCGAGAGCGAGCGCCCGCAGCGCCGAGA 1359  
 Db 22 TGAAGAGGTGGAGAAATGAAGACAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 81  
 QY 1360 TGAAGAGGAGAGAGGTGGCGCTGATGAACCCAGATCCGGGCAATCCAAATCTGGAGACGG 1419  
 Db 82 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTTTTCCTCAAACTGGGAATTTG 135  
 QY 1420 CGCAATGATACGCGAATATCAGAGCACGCTGGAAATTTGTGCCGCTGCTCGATGCCGAGG 1479  
 Db 136 CCGAATGATTAAGAAATTTGGGCTACTTTGGAAATGTCATCCACTTACTATGATGCTATC 195  
 QY 1480 CCGTCGATGACCATCAGATCAGATGTCGTCGTCGCGCAAGCTGCCATTTAGCGCAAGGAGG 1539  
 Db 196 CTATCGAAGAGCAGAGATATGTCGTCTGTAGGAAACCCCACTGTAATGAAGCAAGAT 255  
 QY 1540 TCAATCGCAAGGAGATCGATGTCATTTTCGGTGGCGCGCAAGGACATGCTCATGTGCAAG 1599  
 Db 256 TGGCAAGAAAGAAATGATGTGATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 315  
 QY 1600 AGCGCGCCAGAGGTGCGACTCACCAGATTTCTGGAGAGACCAAGTTTCGCTTCGACT 1659  
 Db 316 AACCAAGTTGAAGTGGAGCTTAAACAAAGTATCTGGAGAACCAAGCAATCTGCTTTGACT 375  
 QY 1660 ACGCCTTCAACGACACGTGCGCAATGCCATGGTATACAAATACACAGCAAGCCGTTGG 1719  
 Db 376 TTGCAATTTGATGAACAGCTTCGAAATGAAGTGTCTACAGTTTCACAGCAAGGCCACTGG 435  
 QY 1720 TGAACACCATTTTCGAGGGCGGAAATGGCGAGCTGCTTTCGCTACGGCAGACGGGATCGG 1779  
 Db 436 TACAGACAATCTTTGAAGGTGGAAAGCACTTTGTTTTCATATATGCGCAGACAGGAAGTG 495



Db 556 TCTATGCCATGGCCCTCCGGGAGCTTCTCTCCTGAAGAATCAACCCCTGCTACCGGAAGT 615  
Qy 1900 TGAATCTAGTCTCTCGGCCAGTCTTCTTTGAGATTACAGTGGCAAGCTCTTCGATCTTC 1959  
Db 616 TGGGCTGGAAGTCTATGTGACATTTCTTCGAGATCTCAATGGGAAGCTGTTGACCTGC 675  
Qy 1960 TGTCCGCAAGCAGAAATCGCGCTCTGGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 676 TCAACAAGAGGCCAAGCTGCGCTGCTGGAGATGGCAGCAACAGTGCAGTGGTGG 735  
Qy 2020 GACTCACCAGAGAGTGTGATGCGCTGCGAGGAGTACTGAAGTCTCATCAGACACGCCA 2079  
Db 736 GCGTCGAGGAGCATCTGTTAACTCTGCTGATGATGTCATCAAGATGATGACATGGGCA 795  
Qy 2080 ATGTGCTCCGCAACATCCGSCCAGAGCTCGGCCCACTCCCAATTCGTCGGTTCGCAACGCCG 2139  
Db 796 GCGCTGCGAGAACCTCTGGGAGACATTTGCCAACTCCCAATTCCTCCGCTCCCAACGCCGT 855  
Qy 2140 TTTTCCAGATGTGTGCTGGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCCTGTTCA 2199  
Db 856 GCTTCCAAATTAATCTTCGAGCTAAGG-----GAGATGATGGCAAGTTCCTTTGG 909  
Qy 2200 TCGATCTCGCGGCAATGAGCGGGCGTGGACACTTCTCTCGGCCGATCGGCAGACGGCTA 2259  
Db 910 TAGATCTGGCAGGGAATGAGCGGCGGACACTTCCAGTGTCTGACCGGCAGACCCGCA 969  
Qy 2260 TGGAGGGTGGCGAGATTAACAATCGCTGCTGGCCCTCAAGGATGCAATTCGTGGTGG 2319  
Db 970 TGGAGGGCGCAGAAATCAACAGAGTCTCTTAGCCCTGAAGGATGCAATCAGGCGCCCTGG 1029  
Qy 2320 GCAAACTGCGGCCACTTGCCTTCCGTGCTCCAACTCACCAGAGTGTGCGCGACT 2379  
Db 1030 GACAGAACAGGCTCACACCCCGTCTCGTGAGAGCAAGCTGACACAGTGTGAGGACT 1089  
Qy 2380 CGTTCATTTGGCGAGAGAGCAAGAGTGCATGATAGCATGATCTCGCCGGGACTTAGCT 2439  
Db 1090 CTTTCATTTGGGGAGAACTCTAGGACTTGCATGATGTCACCGCATCTCACAGGCATAAGCT 1149  
Qy 2440 CTTGCGAGCACGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492  
Db 1150 CTTGTAATATACTTTAAACACCCCTGAGATATGACAGAGGCTCAAGAGCTG 1202

## RESULT 14

AD24085  
ID AD24086 standard; DNA; 2172 BP.  
XX AC AD24086;  
XX DT 09 APR-2002 (first entry)  
XX DE Human mitotic centromere-associated kinesin protein fragment #7 DNA.  
XX KW Human; MCAK; mitotic centromere-associated kinesin; motor domain;  
XX KW cellular proliferation disorder; cancer; hyperplasia; restenosis;  
XX KW cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;  
XX KW arthritis; graft rejection; inflammatory bowel disease; ds.  
XX OS Homo sapiens.

XX Key Location/Qualifiers  
XX FH 1. 2172  
XX CD \$/\*tag= a  
XX FT /product= "MCAK enzyme fragment"  
XX FT

XX US331424-B1.

XX PD 18 DEC-2001.

XX XX 15 JUN-2000; 2000US-00594669.

XX XX 20 APR-1999; 99US-00295612.

XX XX 18 MAY-1999; 99US-00314464.

XX (CYTO-) CYTOKINETICS INC.  
PA Beraud C, Sakowicz R;  
XX WPI; 2002-089075/12.  
DR P-PSDB; AAE14505.  
XX New human MCAK (mitotic centromere-associated kinesin) protein useful in  
XX identifying agents for use in the treatment of cellular proliferation  
XX disorders.  
XX Disclosure; Fig 13; 44pp; English.  
XX The invention relates to human MCAK (mitotic centromere-associated  
XX kinesin) protein, and its fragments that comprise a motor domain and  
XX directly or indirectly produce ADP or phosphate. The MCAK enzyme and its  
XX fragments are used in methods to identify compounds that modulate their  
XX activity. Modulators of MCAK are useful as therapeutic agents for  
XX treating cellular proliferation disorders such as cancer, hyperplasias,  
XX restenosis, cardiac hypertrophy, immune disorders, inflammation,  
XX autoimmune disease, arthritis, graft rejection, inflammatory bowel  
XX disease. The present sequence is a human MCAK fragment encoding DNA  
XX  
SQ Sequence 2172 BP; 619 A; 518 C; 571 G; 464 T; 0 U; 0 Other;

Query Match 14.8%; Score 454.6; DB 6; Length 2172;  
Best Local Similarity 62.6%; Pred. No. 6.2e-106;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
Qy 1300 TGAAGAGGTGGAGGACTGAAGAGAAATCGCAGAAAGCGACGGCCGACAGGCCGAGA 1359  
Db 578 TGAAGAGGTGGAAAAAATGAAGAAACAGCGAGAGAGAGAGGCCCGCAGAACTCTGAAA 637  
Qy 1360 TGAAGAGGAGAGAGTGGCGCTGATGAACACAGGATCGGGCAATCCAAACTGGGAGACGG 1419  
Db 638 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTCTTCCAAACTGGGAAATTTG 691  
Qy 1420 CGCAATGATACGGGAATATCAGAGCAGCGTGGAAATTTGCGCTCTCGATGGCCAGG 1479  
Db 692 CCGGAATGATTAAGAAATTTGCGGGCTACTTTTGGAAATGTCATCCACTTACTATGACTGATC 751  
Qy 1480 CCGTGGATGACCATCAGATCAGATGCGTGGCGAAGCGTCCCATTAGCGCAAGGAGG 1539  
Db 752 CTATCGAGAGACAGAAATATGTCTGTGTAGGAAACGCCCACTGCAATAAGCAAGAT 811  
Qy 1540 TCAATCGCAAGGAGATCGATGTCAATTTGCGTGGCGGCAAGGACATGCTCATCGTGACG 1599  
Db 812 TGGCCAAAGAAAGAAATTTGATTTCCATTCTAGCAAGTGTCTCTCTTGTATCATG 871  
Qy 1600 AGCGCGCAGCAAGTGCACCTCACCAGTTCTCTGGAGAACCAAGTTCGCTTCGACT 1659  
Db 872 AACCCAAAGTTGAAAGTGGACTTAAACAAAGTATCTGGAGAAACCAAGCAATTCGCTTTCGACT 931  
Qy 1660 ACGCTTCAACAGCACACGTGGCAACATGCCATGTTATACAAATACACAGCCAAAGCCGTTGG 1719  
Db 932 TTGCAATTTGATGAAACAGCTTCGAATGAAGTGTCTACAGGTTCAACAGGAGCCACTGG 991  
Qy 1720 TGAATAACCAATTTTGGAGGGCGGAATGGCGAGCTGCTTTCGCTACGGCAGACGGGATCGG 1779  
Db 992 TACAGACAAATCTTTGAAGGTGGAAGAAAGCAACTTTGTTTGCATATATGGCAGACAGGAAGTG 1051  
Qy 1780 GCAAAAGCACACATGGCGGTGAGTTTAAATGGAAGGTGCGAGGACTGCAAGNACGCCA 1839  
Db 1052 GCAAGACACATACTATGGGCGGAGACCTCTCTGGGAAAGCCCAAGATGCATCCAAAGGGA 1111  
Qy 1840 TCTACGCCATGGCGGCAAGGATGTCTTGTGACCCCTGAAATATATCCCGCTTACCGGCCA 1899  
Db 1112 TCTATGCCATGGCTCCCGGAGCTCTTCTCTCTGAAGATCAACCTCTGTACCGGAAGT 1171  
Qy 1900 TGAATCTAGTCTGCTCGGCCAGTTTCTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC 1959  
Db 1172 TGGCCCTTGAAGTCTATGTGACATTTCTTCGAGATCTACAAATGGGAAGCTGTTGACCTGC 1231



QY 1960 TGTCCGACAGAGAACTCGCGCTCTGAGGATGGTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 1232 TCACCAAGAGGCAAGCTCGCGCTGCTGGAGATGGCAAGCAACAGGTGCAAGTGGTGG 1291  
QY 2020 GACTCACCCGAGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2079  
Db 1292 GGTGTCAGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGATGGGCA 1351  
QY 2080 ATGCTCCCGAAATCCGCGCAGAGCTCGGCCCAACTCCAAATTCGTCGCGTTCGACGCCG 2139  
Db 1352 GCSCCTGCAGAACTCTGGSCAGACATTTGCCAACTCCAAATTCCTCCGCTCCACGCT 1411  
QY 2140 TTTTCCAGATTGCTGGCGCCGAGGCTCGACGAAGATCCATGGCAAGTTCTCTGTTCA 2199  
Db 1412 GCTTCCAAAATTATCTTCGAGCTAAAG-----GAGAATGCATGGCAAGTTCTCTTTGG 1465  
QY 2200 TCGATCTGGCGGCAATAGCGGGGCTGGACACTTCTCGCGGATCGGCGAGCGCTA 2259  
Db 1466 TAGATCTGGCAGGAATAGCGAGGCGCGACACTTCCAGTGTGCTGACCGGACGCCGA 1525  
QY 2260 TGGAGGTGCGGAGATTAACAATCGCTGCTGGCCCTCAAGGAGTGCATTGCTGCTGG 2319  
Db 1526 TGGAGGCGCAGAAATCAACAGATCTCTTAGCCCTGAGGAGTGCATCAGGGCCCTGG 1585  
QY 2320 GCAACAGTGGGCCACTTGGCCCTTCGGTGTCTCCAACTCAACCCAGGTGCTGCGGACT 2379  
Db 1586 GACAGAAACAGGCTCACACCCGCTTCGGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1645  
QY 2380 CGTTCAATCGGAGAGAGCAGATGTCATGATGATGATGATGATGATGATGATGATG 2439  
Db 1646 CCTTCAATGGGGAACTCTAGGACTTGCATGATGATGATGATGATGATGATGATGAT 1705  
QY 2440 CTTGCCAGCACAGCTCAACACGCTGCTGATGATGATGATGATGATGATGATGATG 2492  
Db 1706 CTTGTGAATATCTTTAAACACCTGAGATATGCAGACAGGTCAGAGGACTG 1758

RESULT 15  
ADE10094  
ID ADE10094 standard; DNA; 2172 BP.  
XX AC ADE10094;  
XX DT 29-JAN-2004 (first entry)  
XX DE DNA encoding human MCAK enzyme fragment #7.  
XX KW Human; motor protein; mitotic centromere-associated kinesin; MCAK;  
KW microtubule-stimulated ATPase activity; cellular proliferation disorder;  
KW cancer; autoimmune disease; arthritis; inflammatory bowel disease;  
KW keloid; psoriasis; tumour; cytotoxic; immunosuppressive; antiarthritic;  
KW antiinflammatory; gastrointestinal; vulnery; antipsoriatic; ds.  
XX OS Homo sapiens.  
XX FH Key  
XX FT CDS  
XX FT 1. 2172  
XX FT /\*tag= a  
XX PN US638754-B1.  
XX PD 28-OCT-2003.  
XX PF 29-MAR-2002; 2002US-00112432.  
XX PR 28-NOV-2000; 2000US-00724215.  
XX PA (CYTO-) CYTOKINETICS INC.  
XX PI Beraud C, Sakowicz R;  
XX WPI; 2003-842789/78.  
XX DR

DR P-PSDB; ADE10095.  
XX New nucleic acid encoding motor proteins, useful in identifying compounds  
PT for the treatment of cellular proliferation disorders, such as cancer,  
PT autoimmune disease, inflammatory bowel disease, arthritis, keloids and  
PT psoriasis.  
XX Claim 2; Fig 13; 44pp; English.  
XX The present invention relates to the isolation of polynucleotide  
CC sequences encoding fragments of human motor protein, mitotic centromere-  
CC associated kinesin (MCAK). The MCAK enzyme fragments have microtubule-  
CC accumulated ATPase activity. The invention also discloses methods for the  
CC use of the motor protein fragments. The methods and compositions are  
CC useful for high throughput screening systems for identifying compounds  
CC useful in the treatment of cellular proliferation disorders, such as  
CC cancer, autoimmune disease, arthritis, inflammatory bowel disease,  
CC keloids, psoriasis and tumours. The present sequence encodes a human MCAK  
CC enzyme fragment.  
XX Sequence 2172 BP; 619 A; 518 C; 571 G; 464 T; 0 U; 0 Other;  
SQ Query Match 14.8%; Score 454.6; DB 10; Length 2172;  
Best Local Similarity 62.6%; Pred. No. 6.2e-106;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
QY 1300 TGAAGAGGTGGAGGACATGAAAGGAGATCGCGAGAGCGCGCCGACAGCCGAGA 1359  
Db 578 TGAAGGAGTGGAAAAATGAAGACACAGCGAGAGAGAGAGGCCAGAACTCTGAAA 637  
QY 1360 TGAAGAGGAGAGGTGGCGCTGATGAACAGAGATCCGGGCAATCCAACTGGGAGACGG 1419  
Db 638 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTTTCCAACTGGGAATTG 691  
QY 1420 CGCAATGATACCGGAATATCAGACGCTCGAATTTCTGCGCTGCTCGATGGCCAGG 1479  
Db 692 CCGGAATGAATTAAGAAATTCGGGTACTTTGGAATGTCTACCTACTATGACTGATC 751  
QY 1480 CCGTGCATGACCATCAGATCAGATGTGCGTGGCAAGCGTCCCATTCGCGCAAGGAGG 1539  
Db 752 CTATCGAGAGCACAGATATGTCTGTGTAGGAAAGCCCACTGAATAAGCAAGAT 811  
QY 1540 TCAATCGCAGAGAGATCGATGCTATTTGGTGGCGGCAAGGACATGCTCATGTCGACG 1599  
Db 812 TGGCCAAAGAAAGAAATTTGATGTGATTTCCATTTCTAGCAAGTGTCTCTCTTGGTACATG 871  
QY 1600 AGCGCGGACGAGGTGCGACCTCACCAGTTCCTGGAGAACCAAGATTTTCGCTTCGACT 1659  
Db 872 AACCCAAAGTTGAAGTGGACTTTAAACAAAGTATCTGGAGAACCAAGCAATTTCTGCTTTGACT 931  
QY 1660 AGCGCTTCAACGACACGTCGCAATGCCATGTTATACAAATACACAGCAAGCCGTTGG 1719  
Db 932 TTGCATTTGATGAACAGCTTCGATGAAGTTGTCTACAGTTTCACAGCAAGCCCACTGG 991  
QY 1720 TGAACCAATTTTCGAGGCGGAATGGCGACGTGCTTCGCTACGSCCAGAGCGGATCGG 1779  
Db 992 TACAGACAATCTTTGAAGGTGGAAAGCAACTTTGTTTGCATATGSCCAGACAGGAAGTG 1051  
QY 1780 GCACAAACGACACCATGGCGGTGAGTTTAATGGAAGGTGAGGACTGCAAGACGGCA 1839  
Db 1052 GCAGACACATATCTATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCAATCCAAAGGGA 1111  
QY 1840 TCTACGCCATGGCGGCAAGGATGCTTTGTGACCTGACCTGAATATCGCGCTTACCGCGCCA 1899  
Db 1112 TCTATGCCATGGCTCCCGGAGAGTCTTCTCTCTGAAGATCAACCTCTGACCGGAAGT 1171  
QY 1900 TGAATCTAGTCTCGGCCAGTTCTTTTGAATTTACAGTGGCAAGGTCTTCGATCTTC 1959  
Db 1172 TGGGCTTGGAGTCTATGTGACATTTCTCGAGATCTACAATGGGAAGCTGTTTGACCTGC 1231  
QY 1960 TGTCCGACAGACAGAACTCGCGCTCTGGAGGATGTTAAACAGCAAGTGGAGGTGGTGG 2019  
Db 1232 TCACCAAGAGGCAAGCTCGCGCTGCTGGAGATGGCAAGCAAGGTCGAAGTGGTGG 1291

QY	2020	GACTCACCAGAGGTGGTCGATGGCGTCGAGGAGTACTGAAGCTCATCCAGCACGGCA	2079
Db	1292	GGCTGCAGAGGACATCTGGTTAACTCTGCTGATGATGTCTCATCAAGATGATCGACATGGGCA	1351
QY	2080	ATGCTGCCGGAACATCCGGCCAGACGTCCGGCCAACTCCAAATTCGTGGGTTCCGACGCGG	2139
Db	1352	GGCCCTGCAGAACCTCTGGGCAGACATTTGCCCACTCCCAATTCCTCCCGCTCCCAAGCGT	1411
QY	2140	TTTTCCAGATTGTCTGGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCTGTTCA	2199
Db	1412	GCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGAAATGCATGGCAAGTTCTCTTTGG	1465
QY	2200	TCGATCTGGCGGGCAATGAGCGGGCGGTGGACACTTCCTCGGCCGATCGGCAGACGGTA	2259
Db	1466	TAGATCTGGCAGGGAATGAGCGAGGCGCGGACACTTCAGTGCTGACCGGCAGACCCGCA	1525
QY	2260	TGGAGGGTGCAGAGATTAACTGCTGGCCCTCAAGGAGTGCATTCGTGCGTTGG	2319
Db	1526	TGGAGGGCGCAGAAATCAACAAGATCTCTTAGCCCTGAAGGAGTGCATCAGGGCCCTGG	1585
QY	2320	GCARACGTCCGCCCACTTGCCCTTCGGTGTCTCCAACTCACCCAGGTGCTGGCGGACT	2379
Db	1586	GACAGACAAGGCTCACACCCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGGACT	1645
QY	2380	CGTTCATTGGCGAGAGAGCAAGCGTGCATGATAGCCATGATCTCGCCGGGACTTAGCT	2439
Db	1646	CCTTCATTGGGGAGAACTCTAGGACTTGCAATGATTGCCACGATCTCACAGGCATAAGCT	1705
QY	2440	CCTGCGAGCACACGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTG	2492
Db	1706	CCTGTGAATATACTTTAAACACCCCTGAGATATGCAGACAGGGTCAAGGAGCTG	1758

Search completed: May 12, 2006, 21:31:17  
Job time : 1779 secs



QY 181 TTTGTTATCCGTTGCGAGCGCAAGCAAAACGACAGTTGGCATTTGAATTCAGCAAAAC 240  
DB |||||  
181 TTTGTTATCCGTTGCGAGCGCAACGAAACGACAGTTGGCATTTGAATTCAGCAAAAC 240  
QY 241 ACACATCTAACGCCGACCGCAAGCAGCACAACAACACTCGGACACTCGAAAAA 300  
DB |||||  
241 ACACATCTAACGCCGACCGCAAGCAGCACAACAACACTCGGACACTCGAAAAA 300  
QY 301 AAAAAACAGACGCTGTCGGCGACCTCGACAGCAGTTCGGTTCGATTTAGTTGTCATGCG 360  
DB |||||  
301 AAAAAACAGACGCTGTCGGCGACCTCGACAGCAGTTCGGTTCGATTTAGTTGTCATGCG 360  
QY 361 CTTGAAATTCGGTTCGGGCTTAGTTTCCACAAGTTTATCGCTCGTCAAGAAACCAACGAAA 420  
DB |||||  
361 CTTGAAATTCGGTTCGGGCTTAGTTTCCACAAGTTTATCGCTCGTCAAGAAACCAACGAAA 420  
QY 421 TAAATTTATTTTCGACCTTAAATAATCTGACTAAATTTGTTTTGTTTTATGTTATTTAT 480  
DB |||||  
421 TAAATTTATTTTCGACCTTAAATAATCTGACTAAATTTGTTTTGTTTTATGTTATTTAT 480  
QY 481 TAGGCACATTTGACACACCAACGTAAGTTACTACATCTAGCACTTAACGGAACTCCTCT 540  
DB |||||  
481 TAGGCACATTTGACACACCAACGTAAGTTACTACATCTAGCACTTAACGGAACTCCTCT 540  
QY 541 GCAAGCAGTGGAAAGTTGCTGCTCAATCAAGCAGTACTCGGAGTTAACGAGGATAAGCCGG 600  
DB |||||  
541 GCAAGCAGTGGAAAGTTGCTGCTCAATCAAGCAGTACTCGGAGTTAACGAGGATAAGCCGG 600  
QY 601 GAGAAAGAAAGAGATCGGTGGAGAAATAGAGATATACAGTGGAGTCAAGAGAAAGGA 660  
DB |||||  
601 GAGAAAGAAAGAGATCGGTGGAGAAATAGAGATATACAGTGGAGTCAAGAGAAAGGA 660  
QY 661 TCATGGACATGATTACGTTGGGCGAGAGGCTCAGATCAAGGGACGATGGCGCGCTCC 720  
DB |||||  
661 TCATGGACATGATTACGTTGGGCGAGAGGCTCAGATCAAGGGACGATGGCGCGCTCC 720  
QY 721 ACATGGCCGTTGGTGGCTGATCAACCAAGTGGGCAAGTGCATCACAGTCAATGGTACG 780  
DB |||||  
721 ACATGGCCGTTGGTGGCTGATCAACCAAGTGGGCAAGTGCATCACAGTCAATGGTACG 780  
QY 781 AGCGCGCGAAACGAAAGGCGAAGAGGTAGAACTGGAGCCCATACTCACGTTCAATCCGG 840  
DB |||||  
781 AGCGCGCGAAACGAAAGGCGAAGAGGTAGAACTGGAGCCCATACTCACGTTCAATCCGG 840  
QY 841 AGCTAATCCAGATACTCTCGAAGCAGCAGCGCGCCCGGAGCCCAAGAAACAGCCACCG 900  
DB |||||  
841 AGCTAATCCAGATACTCTCGAAGCAGCAGCGCGCCCGGAGCCCAAGAAACAGCCACCG 900  
QY 901 CGCCGATGAACCTCTCGCGTAAATCCACACAATCGGCTATCGGTGGCAATCTCACCAAGCC 960  
DB |||||  
901 CGCCGATGAACCTCTCGCGTAAATCCACACAATCGGCTATCGGTGGCAATCTCACCAAGCC 960  
QY 961 GTATGACCATCGCGGAAACATGCTGAAACAAGATCCAGGAAAGCCAGTTCGCAATC 1020  
DB |||||  
961 GTATGACCATCGCGGAAACATGCTGAAACAAGATCCAGGAAAGCCAGTTCGCAATC 1020  
QY 1021 CGATTTGACAGCAATAGCTGATACAAACAGCACTCCAACTACGCGCGCGGAG 1080  
DB |||||  
1021 CGATTTGACAGCAATAGCTGATACAAACAGCACTCCAACTACGCGCGCGGAG 1080  
QY 1081 GTGTGGGCAACCAACGTCGACGACCACTGGATTACAGCGTCCACGGTACTCGCAAGCTG 1140  
DB |||||  
1081 GTGTGGGCAACCAACGTCGACGACCACTGGATTACAGCGTCCACGGTACTCGCAAGCTG 1140  
QY 1141 CTACCGGCGAGCAGCAGCAAGAGATCGCTCGGCGGTGCTTAATAACATTTGCCCAATC 1200  
DB |||||  
1141 CTACCGGCGAGCAGCAGCAAGAGATCGCTCGGCGGTGCTTAATAACATTTGCCCAATC 1200  
QY 1201 CCAGCGCGGACGAGTCTGCTCGGCGGCAACAGGAGTCCGCACTCGGCGCAACACC 1260  
DB |||||  
1201 CCAGCGCGGACGAGTCTGCTCGGCGGCAACAGGAGTCCGCACTCGGCGCAACACC 1260  
QY 1261 AGGAGCTGGCGGCTAGTAGTACCGCGGCGATCGCACGCAATTGAAGAGGTGGAGCGACTGA 1320

DB |||||  
1261 AGGAGCTGGCGGCTAGTAGTACCGCGGCGATCGCACGCAATGARAGAGGTGGAGCGACTGA 1320  
QY 1321 AGGAGATTCGCGAGAGCGCGCCGACAGGCCGAGATGAAGAGAGAGAGGTGGCGC 1380  
DB |||||  
1321 AGGAGATTCGCGAGAGCGCGCCGACAGGCCGAGATGAAGAGAGAGAGGTGGCGC 1380  
QY 1381 TGATGAACCAAGGATCCGGGCAATCCAACTGGGAGCGGCGCAATGATACGCGCAATATC 1440  
DB |||||  
1381 TGATGAACCAAGGATCCGGGCAATCCAACTGGGAGCGGCGCAATGATACGCGCAATATC 1440  
QY 1441 AGAGCAGCTGGAATTTGTCCGCTGCTCGATGCCAGGCGCTGATGACCATCAGATCA 1500  
DB |||||  
1441 AGAGCAGCTGGAATTTGTCCGCTGCTCGATGCCAGGCGCTGATGACCATCAGATCA 1500  
QY 1501 CAGTGTGCGCGCAAGCGTCCCATTTAGCCGCAAGGAGTCAATCGCAAGGAGATCGATG 1560  
DB |||||  
1501 CAGTGTGCGCGCAAGCGTCCCATTTAGCCGCAAGGAGTCAATCGCAAGGAGATCGATG 1560  
QY 1561 TCATTTTCGTCGCGCAAGGACATGCTCATCGTGACGAGCGCGCAGCAAGGTCCGACC 1620  
DB |||||  
1561 TCATTTTCGTCGCGCGCAAGGACATGCTCATCGTGACGAGCGCGCAGCAAGGTCCGACC 1620  
QY 1621 TCACCAAGTTCTCGGAGAACCAAAAGTTTCGCTTCGACTAGCGCTTCAACGACACGTGCG 1680  
DB |||||  
1621 TCACCAAGTTCTCGGAGAACCAAAAGTTTCGCTTCGACTAGCGCTTCAACGACACGTGCG 1680  
QY 1681 ACAATGCGATGGTATACAAATACACAGCAAGCGTTGGTGAACCAATTTTCAGAGGCG 1740  
DB |||||  
1681 ACAATGCGATGGTATACAAATACACAGCAAGCGTTGGTGAACCAATTTTCAGAGGCG 1740  
QY 1741 GAATGGCGAGTGTCTCGCTACGCGCAGACGGATCGGCGCAAAACGACACACCTGGCG 1800  
DB |||||  
1741 GAATGGCGAGTGTCTCGCTACGCGCAGACGGATCGGCGCAAAACGACACACCTGGCG 1800  
QY 1801 GTGAGTTTAATGGAAGGTGCAGGACTGCAAGAACGGCATCTAGCCATGGCGGCCAAGG 1860  
DB |||||  
1801 GTGAGTTTAATGGAAGGTGCAGGACTGCAAGAACGGCATCTAGCCATGGCGGCCAAGG 1860  
QY 1861 ATGTCTTTGTGACCTGAAATATGCGCGTTACCGCGCATGAATCTAGTCTCGGCGCA 1920  
DB |||||  
1861 ATGTCTTTGTGACCTGAAATATGCGCGTTACCGCGCATGAATCTAGTCTCGGCGCA 1920  
QY 1921 GTTCTTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTCTGTCGCAACGACGAACTGC 1980  
DB |||||  
1921 GTTCTTTTGAGATTTACAGTGGCAAGGTCTTTCGATCTTCTGTCGCAACGACGAACTGC 1980  
QY 1981 GCGTCTCGAGGATGTTAAACAGCAAGTGCAGGTGGTGGGACTCACCGAAGGTGGTGC 2040  
DB |||||  
1981 GCGTCTCGAGGATGTTAAACAGCAAGTGCAGGTGGTGGGACTCACCGAAGGTGGTGC 2040  
QY 2041 ATGCGCTCGAGGAGTACTGAAAGCTCATCCAGACGCGCAATGCTGCCGCAACATCCGCGC 2100  
DB |||||  
2041 ATGCGCTCGAGGAGTACTGAAAGCTCATCCAGACGCGCAATGCTGCCGCAACATCCGCGC 2100  
QY 2101 AGAGCTCGGCCCACTCAATTCGTCGCTTCGCAACGCGCTTCCAGATTTGCTGGCGG 2160  
DB |||||  
2101 AGAGCTCGGCCCACTCAATTCGTCGCTTCGCAACGCGCTTCCAGATTTGCTGGCGG 2160  
QY 2161 CGCAGGCTCGAGCAAGATCCATGGCAAGTTCTCGTTTCATCGATCTCGCGGCAATGAGC 2220  
DB |||||  
2161 CGCAGGCTCGAGCAAGATCCATGGCAAGTTCTCGTTTCATCGATCTCGCGGCAATGAGC 2220  
QY 2221 GGGCGTGGACACTTCTCGCGCGATCGGCAGACGCGTATGGAGGGTCCGAGATTAAACA 2280  
DB |||||  
2221 GGGCGTGGACACTTCTCGCGCGATCGGCAGACGCGTATGGAGGGTCCGAGATTAAACA 2280  
QY 2281 AATCGCTGCTGGCCCTCAAGGATGCAATTCGCTGGTTGGGCAACAGTCCGCCCACTTGC 2340  
DB |||||  
2281 AATCGCTGCTGGCCCTCAAGGATGCAATTCGCTGGTTGGGCAACAGTCCGCCCACTTGC 2340  
QY 2341 CCTTCGCTGCTCCAAACTCACCGAGTGTGCGCGACTCGTTTCATTTGGGCGAGAGCA 2400  
DB |||||

Db 2341 CCTTCGGTCTCCAAACTCACCAGGTGCTCGCGACTGCTTCATTGGGGAAGACA 2400  
Qy 2401 AGACGTGCATGATAGCATGATCTCGCGGACTTAGCTCTCGGACACACCTCAACA 2460  
Db 2401 AGACGTGCATGATAGCATGATCTCGCGGACTTAGCTCTCGGACACACCTCAACA 2460  
Qy 2461 CGCTGGCTATCGGATCGTGTCAAGAGCTGCTGCTCAAGGATATCGTCAAGTTTGGC 2520  
Db 2461 CGCTGGCTATCGGATCGTGTCAAGAGCTGCTGCTCAAGGATATCGTCAAGTTTGGC 2520  
Qy 2521 CTGGCGGCACACCGAGCCCATCGAGATCAGGACGACGAGGAGGAGGAGCTCAACA 2580  
Db 2521 CTGGCGGCACACCGAGCCCATCGAGATCAGGACGACGAGGAGGAGGAGCTCAACA 2580  
Qy 2581 TGTGTCATCGCATCTCGCATCAGCTGATCCCAATTCGATGACACCGGCGAGCTCGA 2640  
Db 2581 TGTGTCATCGCATCTCGCATCAGCTGATCCCAATTCGATGACACCGGCGAGCTCGA 2640  
Qy 2641 ATAATCAGGTGCTCGGCTCTCATCTCGGGGGGCTCATTCAGCAATTAATA 2700  
Db 2641 ATAATCAGGTGCTCGGCTCTCATCTCGGGGGGCTCATTCAGCAATTAATA 2700  
Qy 2701 ACAACAAAGAAACGGAACCGCGCAACATGACCTGCGCCATGCTGAGTTCGTGAGCG 2760  
Db 2701 ACACAAACGAAGAACGGAACCGCGCAACATGACCTGCGCCATGCTGAGTTCGTGAGCG 2760  
Qy 2761 AACACGAGATGTCGACGAGCTGATTTGTCAGACACGAGCCCATCGACGACCTGCGAGCAGA 2820  
Db 2761 AACACGAGATGTCGACGAGCTGATTTGTCAGACACGAGCCCATCGACGACCTGCGAGCAGA 2820  
Qy 2821 CGGAGGAGATGTTGGTGGAGTATCATCGCACCGCTTAATGCCACACTGGAGACTTCCTCG 2880  
Db 2821 CGGAGGAGATGTTGGTGGAGTATCATCGCACCGCTTAATGCCACACTGGAGACTTCCTCG 2880  
Qy 2881 CGGAGTCAAGGCGCTGTACAATCTGACCAACTATGTGACTACGACGAGGACTGTACT 2940  
Db 2881 CGGAGTCAAGGCGCTGTACAATCTGACCAACTATGTGACTACGACGAGGACTGTACT 2940  
Qy 2941 GCAACGCGGCGAGTCGATGTTCTCGAGCTGCTGGACATCGCCATCCAGTCCCGGACA 3000  
Db 2941 GCAACGCGGCGAGTCGATGTTCTCGAGCTGCTGGACATCGCCATCCAGTCCCGGACA 3000  
Qy 3001 TGATGCGCAATATCGGCGCAAGTTGGCCAAAGGAGAGATGCTGTCGACGCTCAATT 3060  
Db 3001 TGATGCGCAATATCGGCGCAAGTTGGCCAAAGGAGAGATGCTGTCGACGCTCAATT 3060  
Qy 3061 CGCCGAATGCGAAGCTTAGT 3081  
Db 3061 CGCCGAATGCGAAGCTTAGT 3081

RESULT 2  
US-11-097-143-1103  
; Sequence 1103, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383

; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1103  
; LENGTH: 3081  
; TYPE: DNA  
; ORGANISM: DROSOPHILA  
US-11-097-143-1103

Query Match 100.0%; Score 3081; DB 10; Length 3081;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAACATAAAAAATTTGTTGCTGACATCTGGTGGTTCGCAAACTATTTCTAGCAGATTTT 60  
Db 1 AAACATAAAAAATTTGTTGCTGACATCTGGTGGTTCGCAAACTATTTCTAGCAGATTTT 60  
Qy 61 GTGATATTTTCTGTTGATCGGTGATTAATCCGCCAGTTTTTTTTTAATGGAAGTGCT 120  
Db 61 GTGATATTTTCTGTTGATCGGTGATTAATCCGCCAGTTTTTTTTTAATGGAAGTGCT 120  
Qy 121 AACACATTTGAGCGGTTCGGAAGATAGCAGGAAGAGCAGCGGCTGCGTTTTTCCCTT 180  
Db 121 AACACATTTGAGCGGTTCGGAAGATAGCAGGAAGAGCAGCGGCTGCGTTTTTCCCTT 180  
Qy 181 TTTGTTATCCGTTCGACGCAACGAAACGACAGTTGGCATTTGAAATTCAGCACAAC 240  
Db 181 TTTGTTATCCGTTCGACGCAACGAAACGACAGTTGGCATTTGAAATTCAGCACAAC 240  
Qy 241 ACACATATTAAGCCGACCCGCAAGCAGCACAACACTGGGACACTCGAAAAAA 300  
Db 241 ACACATATTAAGCCGACCCGCAAGCAGCACAACACTGGGACACTCGAAAAAA 300  
Qy 301 AAAAAACAGACGCTGTCGGCGACCTCGCAAGCAGTTCGGTTTGGATTGTCAATGC 360  
Db 301 AAAAAACAGACGCTGTCGGCGACCTCGCAAGCAGTTCGGTTTGGATTGTCAATGC 360  
Qy 361 CTGGAATTCGGTTTCGGGCTTAGTTTCCCAAGTTTATCGCTCGTCAAGAAACGAAA 420  
Db 361 CTGGAATTCGGTTTCGGGCTTAGTTTCCCAAGTTTATCGCTCGTCAAGAAACGAAA 420  
Qy 421 TAAATTTATTTTCGACCTAAAAATCTGACTAAATTTGTTTGTGTTTATGTTATTATT 480  
Db 421 TAAATTTATTTTCGACCTAAAAATCTGACTAAATTTGTTTGTGTTTATGTTATTATT 480  
Qy 481 TAGGCACATTTTGCACACCAACCTAGTTACTACTACATCTAGCAGTAAACGGAACCTCT 540  
Db 481 TAGGCACATTTTGCACACCAACCTAGTTACTACTACTACTAGCAGTAAACGGAACCTCT 540  
Qy 541 GCAAGCAGTGAAGTTGCTGCTCCATCAAGCAGTACTCGGAGTTAACGCGAGTAAGCCGG 600  
Db 541 GCAAGCAGTGAAGTTGCTGCTCCATCAAGCAGTACTCGGAGTTAACGCGAGTAAGCCGG 600  
Qy 601 GAGAAAGAAAGAGATCGGTGGAGATATAGAGATATACAGGTGGAGTCAAGAGGAAGA 660  
Db 601 GAGAAAGAAAGAGATCGGTGGAGATATAGAGATATACAGGTGGAGTCAAGAGGAAGA 660  
Qy 661 TCATGGACATGATTACCGTGGGCGAGCGCTCAAGATCAAGCGGAGTGGCCCGCTCC 720  
Db 661 TCATGGACATGATTACCGTGGGCGAGCGCTCAAGATCAAGCGGAGTGGCCCGCTCC 720  
Qy 721 ACATGGCGGTGGCGGTGATCAACAGTCCGGCAAGTGCATCACAGTCAATGTTAGTACG 780  
Db 721 ACATGGCGGTGGCGGTGATCAACAGTCCGGCAAGTGCATCACAGTCAATGTTAGTACG 780  
Qy 781 AGCGCGGCAACGAGAGGCGAGGAGTGAAGTGGAGCCCATCTACGCTCAATTCGG 840  
Db 781 AGCGCGGCAACGAGAGGCGAGGAGTGAAGTGGAGCCCATCTACGCTCAATTCGG

Db 781 AGCGCGGCGAAACGAAGGCAAGGAGTAGAACTGGAGCCCATCTCACTCAATCCGG 840  
Qy 841 AGCTAATGCAAGATAGTGTGAAACAGCAGCGCGCGCCCGAGCCCAAGCAAGCCACCG 900  
Db 841 AGCTAATGCAAGATAGTGTGAAACAGCAGCGCGCGCCCGAGCCCAAGCAAGCCACCG 900  
Qy 901 CGCCGATGAACCTCTCGGCTAATCCACACAACTCGGCTATCGGTGGCAATCTCACCAGCC 960  
Db 901 CGCCGATGAACCTCTCGGCTAATCCACACAACTCGGCTATCGGTGGCAATCTCACCAGCC 960  
Qy 961 GTATCACCATTGCGCGGAAACATGCTGAACAGATCCAGAAAGCCAGTCTGATTCCTCAATC 1020  
Db 961 GTATCACCATTGCGCGGAAACATGCTGAACAGATCCAGAAAGCCAGTCTGATTCCTCAATC 1020  
Qy 1021 CCAATTGTGAGCAGCAATAGCTGAATACAAACAGCAATCCAACTACGCGCGCGGAG 1080  
Db 1021 CCAATTGTGAGCAGCAATAGCTGAATACAAACAGCAATCCAACTACGCGCGCGGAG 1080  
Qy 1081 GTGTGGCACCACAACTGCGAGCACTGGAATTAACAGCTCCAGGTCTCGCAAGCTG 1140  
Db 1081 GTGTGGCACCACAACTGCGAGCACTGGAATTAACAGCTCCAGGTCTCGCAAGCTG 1140  
Qy 1141 CTACCGGCGAGCAGCAGCAAGAGTTCGCTCGCGGCTGCTTAATAACATTTGCCCAATC 1200  
Db 1141 CTACCGGCGAGCAGCAGCAAGAGTTCGCTCGCGGCTGCTTAATAACATTTGCCCAATC 1200  
Qy 1201 CCAGCGCGCAGCAGCTGCTGCTCGCGCGCAGCAGAGTCCCACTGCGCGCCACAACC 1260  
Db 1201 CCAGCGCGCAGCAGCTGCTGCTCGCGCGCAGCAGAGTCCCACTGCGCGCCACAACC 1260  
Qy 1261 AGGAGCTGGCGGCGCTAGTACC CGCGGATCGCAGCAATGAAGAGGTGGAGCGACTGA 1320  
Db 1261 AGGAGCTGGCGGCGCTAGTACC CGCGGATCGCAGCAATGAAGAGGTGGAGCGACTGA 1320  
Qy 1321 AGGAGATCGCGAGAGCAGCGCGCCGAGAGCGGATGAAGAGAGAGAGTGGCGC 1380  
Db 1321 AGGAGATCGCGAGAGCAGCGCGCCGAGAGCGGATGAAGAGAGAGAGTGGCGC 1380  
Qy 1381 TGATGAACAGAGATCCCGGCAATCCAACTCGGAGAGCGGCAATGATCGCAATATC 1440  
Db 1381 TGATGAACAGAGATCCCGGCAATCCAACTCGGAGAGCGGCAATGATCGCAATATC 1440  
Qy 1441 AGAGCAGCTGGAATTTGTGCGCTGCTCGATGCGCAGCGCTCGATGACCATCAGATCA 1500  
Db 1441 AGAGCAGCTGGAATTTGTGCGCTGCTCGATGCGCAGCGCTCGATGACCATCAGATCA 1500  
Qy 1501 CAGTGTGCTGCGCAAGGCTCCATTTAGCCGCAAGAGGTCAATCGCAGAGAGATCGATG 1560  
Db 1501 CAGTGTGCTGCGCAAGGCTCCATTTAGCCGCAAGAGGTCAATCGCAGAGAGATCGATG 1560  
Qy 1561 TCATTTTGGTCCGCGCAAGGACATGCTCATCGTGCAAGCGCGCAGCAAGGTTCGACC 1620  
Db 1561 TCATTTTGGTCCGCGCAAGGACATGCTCATCGTGCAAGCGCGCAGCAAGGTTCGACC 1620  
Qy 1621 TCAACCAAGTTCTTGAGAACCAACAAAGTTTCGCTTTCGACTAGCCCTTCAACGACACGTGCG 1680  
Db 1621 TCAACCAAGTTCTTGAGAACCAACAAAGTTTCGCTTTCGACTAGCCCTTCAACGACACGTGCG 1680  
Qy 1681 ACAATGCCATGGTATACAAATACACAGCAAGCGGTTGGTGAAACCAATTTTCAGGGCG 1740  
Db 1681 ACAATGCCATGGTATACAAATACACAGCAAGCGGTTGGTGAAACCAATTTTCAGGGCG 1740  
Qy 1741 GAATGGGACCTGCTTCGCTTACGCGCAGAGGATCGGCGCAAAACCCACACCATGGCG 1800  
Db 1741 GAATGGGACCTGCTTCGCTTACGCGCAGAGGATCGGCGCAAAACCCACACCATGGCG 1800  
Qy 1801 GTGAGTTTAAATGGAAGTGCAGGACTGCAAGAACGGCATCTAGCCATGCGCGCCCAAGG 1860  
Db 1801 GTGAGTTTAAATGGAAGTGCAGGACTGCAAGAACGGCATCTAGCCATGCGCGCCCAAGG 1860  
Qy 1861 ATGCTTTTGTGACCTGAATATGCGGTTTACCGCGCATGAATCTAGTGTCTCGGCCA 1920  
Db 1861 ATGCTTTTGTGACCTGAATATGCGGTTTACCGCGCATGAATCTAGTGTCTCGGCCA 1920

Qy 1921 GTTTCCTTTGAGATTTTACAGTGGCAAGCTCTTCGATCTTCTGTCCGACAGCAGAAACTGC 1980  
Db 1921 GTTTCCTTTGAGATTTTACAGTGGCAAGCTCTTCGATCTTCTGTCCGACAGCAGAAACTGC 1980  
Qy 1981 GCGTCTCGAGGATGTAATAACAGCAAGTGCAGGTGGGACTCACCAGAAAGTGGTGG 2040  
Db 1981 GCGTCTCGAGGATGTAATAACAGCAAGTGCAGGTGGGACTCACCAGAAAGTGGTGG 2040  
Qy 2041 ATGCGCTCGAGGATGTAATAACAGCAAGTGCAGGTGGGACTCACCAGAAAGTGGTGG 2100  
Db 2041 ATGCGCTCGAGGATGTAATAACAGCAAGTGCAGGTGGGACTCACCAGAAAGTGGTGG 2100  
Qy 2101 AGAGCTCGGCGCAACTCTCAATTTCTGCGGTTTCCAGACCGCTTTTCCAGATTTGCTCGGC 2160  
Db 2101 AGAGCTCGGCGCAACTCTCAATTTCTGCGGTTTCCAGACCGCTTTTCCAGATTTGCTCGGC 2160  
Qy 2161 CGCAGGCTCGAGCAAGATCCATGGCAAGTTCTGCTTCATCGATCTGCGCGGCAATGAGC 2220  
Db 2161 CGCAGGCTCGAGCAAGATCCATGGCAAGTTCTGCTTCATCGATCTGCGCGGCAATGAGC 2220  
Qy 2221 GGGGCTGGGACACTTCTCGGCGGATCGGCAGACCGCTATGGAGGTGCGGAGATTAA 2280  
Db 2221 GGGGCTGGGACACTTCTCGGCGGATCGGCAGACCGCTATGGAGGTGCGGAGATTAA 2280  
Qy 2281 AATCGCTGCTGGCGCTCAAGAGTGCATTCGCTGGTGGGCAACAGTCCGCCCACTTC 2340  
Db 2281 AATCGCTGCTGGCGCTCAAGAGTGCATTCGCTGGTGGGCAACAGTCCGCCCACTTC 2340  
Qy 2341 CCTTCCGTGCTCTCAAACTCACCCAGTGTGCGCGACTCGTTTCATTTGGCGAGAAAGCA 2400  
Db 2341 CCTTCCGTGCTCTCAAACTCACCCAGTGTGCGCGACTCGTTTCATTTGGCGAGAAAGCA 2400  
Qy 2401 AGAGCTGATGATGAGCATGATCTGCGCGGAACTTAGCTCTGCGAGCACAACGCTCA 2460  
Db 2401 AGAGCTGATGATGAGCATGATCTGCGCGGAACTTAGCTCTGCGAGCACAACGCTCA 2460  
Qy 2461 CGCTGCTATGCGGATCGTCAAGAGCTGCTGCTCAAGGATTCGTTCAAGGATTCGTTGGC 2520  
Db 2461 CGCTGCTATGCGGATCGTCAAGAGCTGCTGCTCAAGGATTCGTTCAAGGATTCGTTGGC 2520  
Qy 2521 CTGCGCGCAGCAGCGGCTCATCGAGATCAACGAGCAGCAGGAGGAGGAGCTCAACA 2580  
Db 2521 CTGCGCGCAGCAGCGGCTCATCGAGATCAACGAGCAGCAGGAGGAGGAGGAGCTCAACA 2580  
Qy 2581 TGGTGCATCCGCACTCGCATGCAATCCCAATTCGATGCAACCGCGCGGCGAGTCA 2640  
Db 2581 TGGTGCATCCGCACTCGCATGCAATCCCAATTCGATGCAACCGCGCGGCGAGTCA 2640  
Qy 2641 ATAATCAGCTGCTCCGGCTCTCATCTCGGCGGCTGCTTCAACCAATTAATA 2700  
Db 2641 ATAATCAGCTGCTCCGGCTCTCATCTCGGCGGCTGCTTCAACCAATTAATA 2700  
Qy 2701 ACAACCAAGAAACGCGGCAACATGAGACCTGCGCTGAGTTCGCTGAGCG 2760  
Db 2701 ACAACCAAGAAACGCGGCAACATGAGACCTGCGCTGAGTTCGCTGAGCG 2760  
Qy 2761 AACACAGATGTCGAGAGCTGATTTGTCAGCAGCAGGCGCATCGAGACCTCGAGCAG 2820  
Db 2761 AACACAGATGTCGAGAGCTGATTTGTCAGCAGCAGGCGCATCGAGACCTCGAGCAG 2820  
Qy 2821 CGCAGAGATGTTGGTGGAGTATCATCGCAGCAGCTTAAATGCGCACCTTCTCTCG 2880  
Db 2821 CGCAGAGATGTTGGTGGAGTATCATCGCAGCAGCTTAAATGCGCACCTTCTCTCG 2880  
Qy 2881 CCGAGTTCGAGGCGCTGTACAATCTGACCAACTATGTTGAGTACGACCGAGTCTG 2940  
Db 2881 CCGAGTTCGAGGCGCTGTACAATCTGACCAACTATGTTGAGTACGACCGAGTCTG 2940  
Qy 2941 GCAACGCGGCGAGTGCATGTTCTCGAGCTGCTGAGCATCGCCATTCAGTGGCGGACA 3000  
Db 2941 GCAACGCGGCGAGTGCATGTTCTCGAGCTGCTGAGCATTCAGTGGCGGACA 3000



QY 3001 TGATGCCGAATATCGCCCAAGTTGGCCCAAGAGAGATGCTGTGTCAGCTTCAATT 3060  
Db 3001 TGATGCCGAATATCGCCCAAGTTGGCCCAAGAGAGATGCTGTGTCAGCTTCAATT 3060  
QY 3061 CGCCGAATGGCAAGCGTTAGT 3081  
Db 3061 CGCCGAATGGCAAGCGTTAGT 3081

RESULT 3

US-11-097-143-1102  
; Sequence 1102, Application US/11097143  
; Publication No. US2005020858A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; PRIOR FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1102  
; LENGTH: 8043  
; TYPE: DNA  
; ORGANISM: DROSOPHILA  
US-11-097-143-1102

Query Match 50.2%; Score 1547.8; DB 10; Length 8043;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1552; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 156 AGCCAGCGGCTGCCGTTTCTTTTGTATCCGTTGCCAGAGCGCAACGAAACGACA 215  
Db 1599 AGCCAGCGGCTGCCGTTTCTTTTGTATCCGTTGCCAGAGCGCAACGAAACGACA 1658  
QY 216 GTTGGCATTTGAATTCAGCACAAACACATACATACCGCGACCCCGCAAGCAGCACAC 275  
Db 1659 GTTGGCATTTGAATTCAGCACAAACACATACATACCGCGACCCCGCAAGCAGCACAC 1718  
QY 276 ACACACACTGGACACTCGAAAARAAAAAAGAGAGCGTGTCCGGGACCTCGACAGGCG 335  
Db 1719 ACACACACTGGACACTCGAAAARAAAAAAGAGAGCGTGTCCGGGACCTCGACAGGCG 1778  
QY 336 TTGGGTTTCGATTTAGTTGTCATGCTTGAATTTCCGTTTCCGGGCTTTAGTTTCCACAGTT 395  
Db 1779 TTGGGTTTCGATTTAGTTGTCATGCTTGAATTTCCGTTTCCGGGCTTTAGTTTCCACAGTT 1838  
QY 396 TATCGCTCGTCAAGAAACCAAGAAATTAATTTTCGACCTAAAAAATCTGACTAAAT 455  
Db 1839 TATCGCTCGTCAAGAAACCAAGAAATTAATTTTCGACCTAAAAAATCTGACTAAAT 1898  
QY 456 TGTGTTTTTGTATGATATTTATTTAGGCATTTTGCACACCAAGCTAGTTACTAC 515  
Db 1899 TGTGTTTTTGTATGATATTTATTTAGGCATTTTGCACACCAAGCTAGTTACTAC 1958

QY 516 ATCTACGACTAACGGAACTCTCTCTGCAAGCAGTGGAAAGTTGCTGTCCATCAAGCAGTAC 575  
Db 1959 ATCTACGACTAACGGAACTCTCTCTGCAAGCAGTGGAAAGTTGCTGTCCATCAAGCAGTAC 2018  
QY 576 TCGGAGTTAAACGAGGATTAAGCCGGGAGAAAGAGAAAGATCGGTGGAGAAATAGAGATA 635  
Db 2019 TCGGAGTTAAACGAGGATTAAGCCGGGAGAAAGAGAAAGATCGGTGGAGAAATAGAGATA 2078  
QY 636 TACAGGTGAGTCAAGAGGAAAGGATCATGGACATGATTTACGTTGGGCGAGCGTCAAG 695  
Db 2079 TACAGGTGAGTCAAGAGGAAAGGATCATGGACATGATTTACGTTGGGCGAGAGCGTCAAG 2138  
QY 696 ATCAAGCGGACGCGCTCCACATGGCGCGTGGTGGCGGTGATCAACCAAGTGGGC 755  
Db 2139 ATCAAGCGGACGCGCTCCACATGGCGCGTGGTGGCGGTGATCAACCAAGTGGGC 2198  
QY 756 AAGTGCATCAGAGTGAATGTGACAGCGCGCGGAAACGAAAGGCAAGAGGTAGAACTG 815  
Db 2199 AAGTGCATCAGAGTGAATGTGACAGCGCGCGGAAACGAAAGGCAAGAGGTAGAACTG 2258  
QY 816 GAGCGCATACTCAGCTCAATCCGGAGCTTAATGCAAGTACTGTGCAAGCAGCAGCGCGC 875  
Db 2259 GAGCGCATACTCAGCTCAATCCGGAGCTTAATGCAAGTACTGTGCAAGCAGCAGCGCGC 2318  
QY 876 CCGGAGCCCAAGAAACCAAGCCACCGCGCGATGAACCTCTCGCTAATCCCAACCAATCG 935  
Db 2319 CCGGAGCCCAAGAAACCAAGCCACCGCGCGATGAACCTCTCGCTAATCCCAACCAATCG 2378  
QY 936 GCTATCGGTGGCAATCTCAAGCGGTATGACCATGGCGGAAACATGCTGAACCAAGATC 995  
Db 2379 GCTATCGGTGGCAATCTCAAGCGGTATGACCATGGCGGAAACATGCTGAACCAAGATC 2438  
QY 996 CAGGAAAGCCAGTCCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1055  
Db 2439 CAGGAAAGCCAGTCCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2498  
QY 1056 AACTCCAACTACCGCGCGGAGGTGTGTGACCAACCAAGCTGACGACCACTGGATTA 1115  
Db 2499 AACTCCAACTACCGCGCGGAGGTGTGTGACCAACCAAGCTGACGACCACTGGATTA 2558  
QY 1116 CAGCGTCAAGTACTCGCAAGCTGTACCGCGCAGCAGCAGCAAGGATTCGCTCGCG 1175  
Db 2559 CAGCGTCAAGTACTCGCAAGCTGTACCGCGCAGCAGCAGCAAGGATTCGCTCGCG 2618  
QY 1176 GTGCTTAATACACATTCGCAATCCAGCGCGCAGCAGCTGTGTTCCGGGCGACAA 1235  
Db 2619 GTGCTTAATACACATTCGCAATCCAGCGCGCAGCAGCTGTGTTCCGGGCGACAA 2678  
QY 1236 GGAGTCGCGCACTCGCGCCCAACCAAGCGGAGCTGGCGCGCTAGTACCGCGCGATCGCAC 1295  
Db 2679 GGAGTCGCGCACTCGCGCCCAACCAAGCGGAGCTGGCGCGCTAGTACCGCGCGATCGCAC 2738  
QY 1296 GCATTGAAGAGGTGGAGCACTGAAGGAGAAATCGCGAGAGCAGCGCGCCGACAGGCGC 1355  
Db 2739 GCATTGAAGAGGTGGAGCACTGAAGGAGAAATCGCGAGAGCAGCGCGCCGACAGGCGC 2798  
QY 1356 GAGATGAAGGAGGAGAGGTGGCGCTGATGAACCAAGGATTCGGGCAATCCAACTGGGAG 1415  
Db 2799 GAGATGAAGGAGGAGAGGTGGCGCTGATGAACCAAGGATTCGGGCAATCCAACTGGGAG 2858  
QY 1416 ACGGCGCAATATGATACGCGAAATATCAGAGCAGCTGGAAATTTGTCGCGTGTCCATGCG 1475  
Db 2859 ACGGCGCAATATGATACGCGAAATATCAGAGCAGCTGGAAATTTGTCGCGTGTCCATGCG 2918  
QY 1476 CAGCGCGTGCATGACCATCAGATCAGTGTGCGTGGCGCAAGCGTCCCATAGCCGCAAG 1535  
Db 2919 CAGCGCGTGCATGACCATCAGATCAGTGTGCGTGGCGCAAGCGTCCCATAGCCGCAAG 2978  
QY 1536 GAGTCAATCCCAAGGAGATCGATGTGATTTGCGTGGCGCGCAAGGATGCTCATCGTG 1595  
Db 2979 GAGTCAATCCCAAGGAGATCGATGTGATTTGCGTGGCGCGCAAGGATGCTCATCGTG 3038

Qy	1596	CACGAGCGGCGACGAAAGTTCGACTCACCAAGTTCTCTGGAGAACCAAGATTTTCGGTTC	1655
Db	3039	CACGAGCGGCGACGAAAGTTCGACTCACCAAGTTCTCTGGAGAACCAAGATTTTCGGTTC	3098
Qy	1656	GACTACGCGCTTTCAACGCAACAGTGGCGCAATGCCATGGTATACAATATACACAGCCAAAGCC	1714
Db	3099	GACTACGCGCTTTCAACGCAACAGTGGCGCAATGCCATGGTATACAAGTCAGATATCAAAATCC	3157

## RESULT 4

```

US-11-097-143-5291
; Sequence 5291, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 5291
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-5291

```

Query Match	18.3%;	Score 564.6;	DB 10;	Length 2190;
Best Local Similarity	68.1%;	Pred. No. 1e-157;		
Matches 834;	Conservative 0;	Mismatches 379;	Indels 12;	Gaps 3;
QY	1296	GCATTGAAAGGTTGGAGCGCACTGMAGGAGAAATCGGAGAAAGCGACGCGCCCGACAGGCC	1355	
DB	493	GTAGTGGGAGGTGAATCGCATGAAGGACAGAGGGAGAGCGAAGGGCTCGCCAGGCG	552	
QY	1356	GAGATGAAGGAGGAGAAAGTGGCGCTGATCAACACAGGATCCGGGCAATCAAACTGGGAG	1415	
DB	553	GAACAGCTCAGAGAAAGATGACTGCTGCGCAATATCCGGGGATCCCACTGGGAG	612	
QY	1416	ACGGCGCAATGATACGCGAATATCAGAGCACGCTGGAATTTGTGCGGCT-----GCTC	1469	
DB	613	GTGTGGTGTGCTGCGCCAATACCGCTCCACCTTGATCTTTTCTCCACTTCGATGCCCTG	672	
QY	1470	GATGGCCAGGCGCTCGATGACCATCAGATCACAGTGTGCGTGCAGACGGTCCCATTAGC	1529	
DB	673	GATCCCAATGGAGGCACTGTTTCAGCAAAATACCGTGTGTGTGCGAAACAGCCCATGAGT	732	
QY	1530	CGCAAGGAGTCAATCGCAAGGAGATCGATGTGTAATTTTCGGTGGCGGCAAGGACATGCTC	1589	
DB	733	CGCAAGGAGAGAACTTCAAGAACCTTGGACATATACAGTTCCAGTGCCTGACGCTG	792	
QY	1590	ATCGTGCAAGCGCGCGACAGAGGTGCACTCACCAAGTTCTCTGGAGAACCAACAAGTTT	1649	
DB	793	ATCGTCAATGATTTGGCGCTCAAGGTGGATCTCACCAAGTTCTCTGGAGAACCAACAATTC	852	

## RESULT 5

```

RESULI 3
US-11-097-143-5290
; Sequence 5290, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: C1.000728

```

Qy	1650	CGCTTCGACTACGCTTCCTCAACGACACGCTGGCA	CAATGCGCAATGGTATACAAATACACAGCC	1709
Db	853	CGTTTCGACTACACGTTTCGACGAGGAGTGCTCCAA	TGCGTGTGTTCTACGATCACACTGTCT	912
Qy	1710	AAGCCGTTGTGTAAACCAATTTTCGAGGGCGGAAT	GGCGACGTGCTTCCCTACGCGCCAG	1769
Db	913	CGTCCGTTGATCAGAACCATGTTTCGAGGGCGGCA	ATGCCACTTGTTTCGCTTACGGACAA	972
Qy	1770	ACGGGATCGGGCAAAACGCAACACATCGGCGGT	GAAGTTTAAATGGAAGGTGCAGGACTGC	1829
Db	973	ACTGGCAGCGGAAAAACGACACATCGGGCGAGAA	TTCTTCGAAAGGTTCAAGGATTGC	1032
Qy	1830	AAGAAAGGCATCTACGCCATCGGCGGCCAAGGA	TGCTTTGTGCACCTGAAATATGCCCGT	1889
Db	1033	GGTACCGGGATCTACGCCATCGGCAGCTCGCGA	TGTCTTCGAGGAGGTATCGCGCCCGGAG	1099
Qy	1890	TACCGGCCATGAATCTAGTCTGCTCGGCCAGT	TTCTTTTCAGATTTTACAGTGGCAAGTCT	1949
Db	1093	TACGGCAATGGTGCCAGATTACGTGCAGCTTCT	TTCGAAATCTATGACCAAGGTG	1152
Qy	1950	TTTGATC---TTCTGTTCGCAACAGAGAAAC	TCTCGCGTCTCTGGAGATGGTAAACAGCAA	2006
Db	1153	TTCGATCTCTTGTATCCCAACAAAGCCCATCT	CTCGGGTCTTAGAGATGTCGAGGCAGCAG	1212
Qy	2007	GTGCAGGTGTGGGACTCACCGAGAAGTGGT	TCGATGGCGTTCGAGCAGGTACTTGAAGTCT	2066
Db	1213	GTGCTGTGTGGGCTTACCGAGATGCGG	TGACCAAAAGTGGAGGATGTCTCTGACATG	1272
Qy	2067	ATCCAGACGGCAATGCTGCCGAAACATCCGG	CCAGACGTCGCGCCAACTCCAATTCGTCTG	2126
Db	1273	ATTGAGCACGGCAGCAAGAGCGCACTTCGG	CCAAACATCGCGCAACGCCAAGTCATCG	1332
Qy	2127	CGTTCGCACGCCGTTTTCAGATGTGCTGCG	CGCCGCGCAGGCGTCGACGAGATCCATGCG	2186
Db	1333	CGTTCCACGCCGTCTTTCAAATAGCACT	CCACTTTCCCGATTTCTGTGGGGCCCAACGGC	1392
Qy	2187	AAGTTCTCTGTTTCATCGATCTGGCGGGCA	ATGACGGGGCGGTGGACACTTCTCTCGGCCGAT	2246
Db	1393	AAGTGTCTCTTCTGTGCACTTGGCGGCA	ATGAACGGGGCGGATACGCAATCTCGCCGAT	1452
Qy	2247	CGGCAGACGGTATGAGGGTGCAGGATTA	CAAAATCTGCTGTGGCCCTCAAGNGTGC	2306
Db	1453	CGTCAAATCTCGCATCAGGGAGCCGAGAT	CAATAAATCTCTGTGGCCCTCAAGGAGTGC	1512
Qy	2307	ATTCTGTGCGTTGGGCAAAACAGTCGGG	CCACTTTCGCCCTTCGCTCTTCCAAAATCA	2366
Db	1513	ATTTCGAGCCCTCAGCGGCGAGTCGAGT	CACTTCCCTTCGTTGGCTTCCAAGTTGACCCAA	1572
Qy	2367	GTGCTGGCGACTCGTTTCATTTGGCG---	AGAGAGCAAGAGTGTGCATGTATAGCCATGATC	2423
Db	1573	GTGCTGGCGACTCTCTTGTTCGGCGGCA	AGAAACAGACCTCTGATGTATGTCATGATA	1632
Qy	2424	TCCCGGGACTTACGTCCTTCGGAGCACG	CTCAACACGCTCGCTTATCGGATCTGCTGTC	2483
Db	1633	TCGCCATCCATGAGCTGCGTGGAGAA	TACGCTCAACACTCTTACGTTACGCACACAGGGTT	1692
Qy	2484	AAGGAGCTGGTGTCAAGGATATCG		2508
Db	1693	AAGGAGCTCATAGCCAAGGAAGCG		1717

1 APPLICANT: UC SA:  
 2 TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
 3 TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
 4 TITLE OF INVENTION: DROSOPHILA GENES.  
 5 FILE REFERENCE: CL000728

```

: CURRENT APPLICATION NUMBER: US/11/097,143
:
: CURRENT FILING DATE: 2005-04-04
:
: PRIOR APPLICATION NUMBER: 60/157,832
:
: PRIOR FILING DATE: 1999-10-05
:
: PRIOR APPLICATION NUMBER: 60/160,191
:
: PRIOR FILING DATE: 1999-10-19
:
: PRIOR APPLICATION NUMBER: 60/161,932
:
: PRIOR FILING DATE: 1999-10-28
:
: PRIOR APPLICATION NUMBER: 60/164,769
:
: PRIOR FILING DATE: 1999-11-12
:
: PRIOR APPLICATION NUMBER: 60/173,383
:
: PRIOR FILING DATE: 1999-12-28
:
: PRIOR APPLICATION NUMBER: 60/175,693
:
: PRIOR FILING DATE: 2000-01-12
:
: PRIOR APPLICATION NUMBER: 60/184,831
:
: PRIOR FILING DATE: 2000-02-24
:
: PRIOR APPLICATION NUMBER: 60/191,637
:
: PRIOR FILING DATE: 2000-03-23
:
: NUMBER OF SEQ ID NOS: 43008
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 5290
:
: LENGTH: 4190
:
: TYPE: DNA
:
: ORGANISM: DROSOPHILA
:
: US-11-097-143-5290

```

Query Match 18.3%; Score 564.6; DB\_10; Length 4190;

1296	QY	GCATTTGAAGAGGTGGAGCGACTGAAGGAGATCGCGAAGACGCGCCCGACAGGCC	1355
1493	Db	GTAGTGGGGAGGTGAATCGCATGAAGGAGCAGAGGCGAAGCAAGGGCTCGCCAGCGC	1552
1356	QY	GAGATGAAGCAGAGAGAGGTGGCGCTGATGAAACGAGGATCGGGGCATCCAAACTGGGAG	1415
1553	Db	GAACAGCTCCAGAGAGAGGATGCACCTCGCTGCGCAATAATCCGGGGAATCCCAACTGGGAG	1612
1416	QY	ACGGCGCAATGATACGCGAATATCAGAGCACGCTGGAAATTGTGTCCGCT-----GCTC	1469
1613	Db	GTGTCGGTGATGTGGCCCAATACCGCTCCACCTTGATCTTTTCTCCACTTCGATGCCCTG	1672
1470	QY	GATGGCCAGCGCGTCGATGACCATCAGATCACAGTGTGCGTGCGCAAGCGTCCCATTAGC	1529
1673	Db	GATCCCAATGGAGGCACTGTTTCAGCAAAATTACGGTGTGTGCGGAAACGACCCATGAGT	1732
1530	QY	CGCAAGGAGGTCAATCGCAAGGAGATCCGATGTCAATTCGGTGC CGCGCAAGGACATGCTC	1589
1733	Db	CGCAAGGAGGAGAACTCCAAGACCTGGACATCATCACAGTTCACAGTGC CGACAGCCTG	1792
1590	QY	ATCGTGACGAGCCGCGCAGCAAGTTCGACCTCACCAAGTTCTCTGGAGAAACCAAGTTT	1649
1793	Db	ATCGTCCATGAGTTGGCGCTCAAGAGTGGATCTCACCAAGTTCTCTGGAGCACCACAAATTC	1852
1650	QY	CGCTTCGACTACGCCCTTCAACGACACGTGGCGCAATGCCCATGTATACAATACACAGCC	1709
1853	Db	CGTTTCGACTACAGTTTCGACGAGGAGTGTCCAAATGCGCTGGTCTACGATCACTGCT	1912
1710	QY	AAGCGTTGGTGA AAAACCAATTTTCGAGGGCGGAAATGCGCAGCTGCTTCGCTACGGCCAG	1769
1913	Db	CGTCCGTTGATCAGAAACCAATGTTTCGAGGGCGGCAATGCCACTTGTTCGCTTACGGACAA	1972
1770	QY	ACGGGATCGGGCAAAAACGACACCATGGGCGGTGAGTTTAATCGGAAGGTGAGGACTGC	1829
1973	Db	ACTGGCAGCGGAAAAACGACACATGGGCGGAGAAATCTTTCGGAAGGTTCAAGATTCG	2032
1830	QY	AAGAACGGCATCTACGCCATCGCGCGCCAAAGATGTCTTTGTGACCTGAAATATGCCGCGT	1889
2033	Db	GGTACCGGATCTACGCCATGCGCAGCTCGCGATGTCTTCGAGGAGGTATCGCGCCGGAG	2092
1890	QY	TACCGCGCCANTGAATCTAGTGTCTCGGCCAGTTTCTTTGAGATTTTACAGTGCAAGGTC	1949
2093	Db	TACCGGCAATGGGTGCCAAGATTACGTGCAGCTTCTTCGAAATCTATGACCAACCAAGGTG	2152

```
; SEQ ID NO 5990
; LENGTH: 1970
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-5990

Query Match      16.4%; Score 505.8; DB 10; Length 1970;
Best Local Similarity 64.7%; Pred. No. 4.3e-140;
Matches 822; Conservative 0; Mismatches 432; Indels 17; Gaps 4;

QY 1252 CCACAAACCAGGAGTGGCGGCTAGTACCAGGATCGCAGCATTTGAAAGAGTGG 1311
Db 399 CGGCCACGACAGACGCGCGCGCAGTTCGAGAGGA--CGAGTGGTGCACACAGGCTG 456
QY 1312 AGCGACTGAAGAGAGTCCGAGAGAGCGCGCCGACAGCGCGAGATGAAGAGGAGA 1371
Db 457 AAGAAATGCGAAAGAACCGGAGCGGAGAGAGACCAAGCCAGGACTCGTCTAGATC 516
QY 1372 AGGTGGCGCTGATGAACCAAGGATCCGGGCAATCCAAACTGGGAGACGGCGCAAAATGATAC 1431
Db 517 GGGAGCAGGGGAAGAACGAGATCCGGGAATCCCACTGGGAAGTAGCCAGATGATAC 576
QY 1432 GCGAATATCAGACGCTGGAATTTGTGCGCTGCTGATGG-----CCAGGCGG 1482
Db 577 GACTGCAACGCGAGCAAAATGGAGAGTCAAGCGGTGAGAAAGTGTACTACGAACGAACGAA 636
QY 1483 TCGATGACCATCAGATCAGATGTCGTCGCGCAAGCGTCCCATTTAGCGCGAAGAGGTCA 1542
Db 637 TCAATTGCCACCAATATTGTTTGTGAGGAGAGACCACTGAGGCGCAGAGAGTGG 696
QY 1543 ATGCCAAGGAGATCGATTCATTTCCGTCGCCGCAAGGACATGCTCATCGTGACAGAC 1602
Db 697 CTGACCGGGAAACAGGATGGTGCAGCATTCGCTCTAAGCACACATTTGGTGGTCCAGAGC 756
QY 1603 CGCGCAGCAGAGTGCAGCTCACAAGTTCCTCGGAGAACACAAAGTTTCGTTGCACTAG 1662
Db 757 CCGCAAGCATGGAACCTGGTCAAGTTCTCTGAAAATCATAGCTTCGGTTTCGATTACG 816
QY 1663 CTTTCAACGACACGTCGACATGTCATGTTATACAAATACACACGCAAGCGTGGTGA 1722
Db 817 TCTTCGACGAGAGTGCTCAATGCCAGGTCTAGAAATTCACGCCGACCTTGATAA 876
QY 1723 AAACCATTTTCGAGCGCGAATGGGAGTGTCTCGCCTACGGCCAGACGGGATCGGCA 1782
Db 877 AGCACAATTTTATGCGGGAATGGCCAGTGTTCGCTACGCAACAATCGAAGCGCA 936
QY 1783 AAACCCACACATGGGCGGTGATTTAATGAAAGGTGAGGACATGCAAGAACGCGCATCT 1842
Db 937 AGACCTATACGATGGGTGTGAGTTCCCGGAAGGCATCAGAGCTCAATGGATGGCATCT 996
QY 1843 AGCCCATGGCGGCAAGGATGTCCTTGTGACCCCTCAATATCGCGTTACCGGCCATGA 1902
Db 997 ATGCAATGGCGCTAAGGACGTTCTCCACTCTAAGACGGTTCCTTATACAGCTTA 1056
QY 1903 ATCTAGTCTCTCGGCCAGATTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC--- 1959
Db 1057 ATCTGAAAGTTTACTGCGAGCTTCTTCGAGATCTACGGCACCGGGTTCGATCTCTGA 1116
QY 1960 TGTCGACACAGCAAAATCGCGCTCTGGAGGATGGTAAACAGCAAGTGCAGAGTGGTGG 2019
Db 1117 TGCTTGGGAAGCCACAACTGCGGTGCTTTGGAGGATAGAAACAGCAGGTGCAAGTGGTGG 1176
QY 2020 GACTCACCGAAGAGTGTTCGATGCGTCGAGGAGTACTGAAGCTCATCCAGCACGCA 2079
Db 1177 GCCTCACCGAATCCAGTACGAACACCGCGGAAGTTCTGGAACCTACTCGAGTTGGGCA 1236
QY 2080 ATGTGCGCCGAACATCCGCGCAGAGCTCGGCCAACTCCAATTCGTTCGGTTTCGACACGCG 2139
Db 1237 ATAGTGTCCGAACCTCGGCTCACACCTCTGCGCAATTCGAAGTCTCTCCGATCGCATGCTG 1296
QY 2140 TTTTCCAGATGTGCTGGCGGCGCAGGCTCGACGAAGATCGATGGCAAGTTCTCGTTCA 2199
Db 1297 TGTTCCAAATCGTGTGAGATCCGCGGCGGCGAGAGCTACACGGGAAATTTCTCGCTTA 1356

RESULT 7
US-11-097-143-5989/c
; Sequence 5989, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5989
; LENGTH: 3970
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-5989

Query Match      16.4%; Score 505.8; DB 10; Length 3970;
Best Local Similarity 64.7%; Pred. No. 6.5e-140;
Matches 822; Conservative 0; Mismatches 432; Indels 17; Gaps 4;

QY 1252 CCACAAACCAGGAGTGGCGGCTAGTACCAGGATCGCAGCATTTGAAAGAGTGG 1311
Db 2572 CCGCACCGCAGCAGACGGGCGCGCAGTTCGAGAGGA--CGAGTGGTGCACACAGGCTG 2515
QY 1312 AGCGACTGAAGAGAGTTCGAGAGAGCGCGCCGACAGGCGCGAGATGAAGAGGAGA 1371
```

Db 2514 AAGAATGCCAAGAACGCGGAGCGAGGCGAGAGCACAAGCCAGGACTCGTCTAGATC 2455  
Qy 1372 AGTGGCGCTGATGAACACAGGATCCGGGCAATCCAACTGGGAGACGGCGCAATGATAC 1431  
Db 2454 GGGAGCAGGGGGAAGAACAGAGATCCGGGAAATCCCAACTGGGAAGTAGCCAGAATGATAC 2395  
Qy 1432 GCGAATATCAGACGACGCTCGGAATTTGTGCGGCTGCTCGATGG-----CCAGGCCG 1482  
Db 2394 GACTGCAACCGGAGCAAAATGGAGAGTCAGCGGGTGAGAAGTGGTACTACGAACGAACGAA 2335  
Qy 1483 TCGATGACCATCAGATCACAAGTGTGCGTGGCGCAAGCGTCCCATTTAGCCGCAAGGAGTCA 1542  
Db 2334 TCAATTGCCACCAATATATGTTGTGTGAGGAAGAGACCACTGAGCGCGNAGSACTGG 2275  
Qy 1543 ATCGAAGAGATCGATGTCATTTTCGTGCGCGCAAGGACATGCTCATGTCACAGAC 1602  
Db 2274 CTGACCGGGAACAGGATGTGTCAGCATTCGCTTAAGCACACATTTGGTGGTCCACGAGC 2215  
Qy 1603 CGCGCAGCAGTCCGACCTCACCAAGTTCCTGAGAACCAAGATTTTCGCTTCGACTACG 1662  
Db 2214 CCGCAAGCATGTGAACCTGGTCAAGTTCCTGGAAATCATAGCTTCGCTTCGATTACG 2155  
Qy 1663 CTTTCAACGACAGCTGCGCAATGCCATGGTATACAAATACACACCAAGCCGTTGGTGA 1722  
Db 2154 TCTTCGACGAGGATGCTCCATGCCACGCTCTACGAATTCACAGCCGACCCCTTGATTA 2095  
Qy 1723 AAACCAATTTTCGAGGCGGGAATGGCAGCTGCTTCGCTACGCGCAGACGGGATCGGCA 1782  
Db 2094 AGCACATTTTGTATGCGGAATGGCCACGCTGTTTCGCCCTACGCAACAACTGGAAGCGCA 2035  
Qy 1783 AAACGACACCATGGGCGGTGAGTTTAAATGGAAGGTGAGAGCTCAGAGACCGGATCT 1842  
Db 2034 AGACCTATACGATGGGTGGTCAAGTTCCTCCGGAAGGCATCAGAGCTCAATGGATGGCATCT 1975  
Qy 1843 ACGCCATGGCGGCCAAGGATGCTTTGTGACCTGAATATGCGCGGTACCGCGCCATCA 1902  
Db 1974 ATGCAATGGCGGTAAAGACGCTGTTCTCCACTTAAAGACGGTTCCTATTAACAGCTTA 1915  
Qy 1903 ATCTAGTCTCTCGGCCAGTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC--- 1959  
Db 1914 ATCTGAAAGTTTACTCGAGCTTCTTCGAGATCTACGCGCACCCGGGTGTTTCGATCTCTCA 1855  
Qy 1960 TGTCCACACAGCAGAACTGCGCGTCTCGAGATGTTAAACAGCAAGTGCAGGTGGTG 2019  
Db 1854 TGCTGGCAAGCCAACTGCGTGTCTTGAGGATAGAAAACAGCAGGTGCAAGTGGTG 1795  
Qy 2020 GACTCACCGAGAAGTGTGATGGCTGAGAGGATCTGAAGCTCATCCAGCAGCGCA 2079  
Db 1794 GCCTACCCAGAAATCCAGTACAGAACACCGCCGAGGTTCTGGACCTACTCGAGTTGGGCA 1735  
Qy 2080 ATGCTCCCGAACATCCGGCCAGACGTCGGCCAACTCCAAATGCTGCGCTTCGACGCGC 2139  
Db 1734 ATAGTCTCCGAACTCTGGGTACACCTCTGCGCAATTCGAACTCTCCGATCGCATGCTG 1675  
Qy 2140 TTTTCCAGATTTGTCTGGCGCGCAGGGCTCGACGAAGATCCATGGCAAGTTCTCGTTCA 2199  
Db 1674 TGTTCAAAATCGTGTGAGATCCGCGCGGCGAGAAAGCTACACGGGAATTTCTCGCTTA 1615  
Qy 2200 TCGATCTGGCGGCATGAGCGGGGTGAGCACTTCCTCGCCGATCGGCGAGACGCTA 2259  
Db 1614 TAGATCTGGCGGGAATGAAAGAGGAGCGGACAAACAGCTCGCGGATCGACACGCGCC 1555  
Qy 2260 TGGAGGTTCGAGATTAAACAAATCGCTGCTGCGCCCTCAAGAGTGCATTTCTGTGGTTGG 2319  
Db 1554 TGGAGGATCCGAGATCAATAATCGCTGCTGTTCTTCAAGGATGCAATTCGCGCTCTGG 1495  
Qy 2320 GMAACAGTCCGCCCACTTGCCCTTCCGTTCTCCAAACTCACCCAGGTGCTGCGGACT 2379  
Db 1494 GCGCCAGTCCGATCAATTTGCCATTCGCTGGTCCAAAGCTGACCCAAAGTCTCGCGGACT 1435  
Qy 2380 CGTTCAATTCGCG---AGAGAGCAGACGTCGATGATGACCATGATCTCGCGGACTTA 2436  
Db 1434 CTTTACGAGGTAAAGAGGTGAAACCTTGATGATGCCATGATCTCGCCATGCTTGC 1375

Qy 2437 GCTCTCGAGCACACGCTCAACACGCTGCGCTATCGCGATCGTGTCAAGGAGCTGGTGG 2496  
Db 1374 ATTGGTGGAGCATACCTTTGAACAGCTGCGTTATCGCGATCGGTGGAAGGAACTAAGTG 1315  
Qy 2497 TCNAGGATATC 2507  
Db 1314 TGGAGTCGATC 1304

RESULT 8  
US-10-651-510-1  
; Sequence 1, Application US/10651510  
; Publication No. US20040096949A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042A  
; CURRENT APPLICATION NUMBER: US/10/651,510  
; PRIOR FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US/09/724,215  
; PRIOR FILING DATE: 2008-11-28  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; PRIOR APPLICATION NUMBER: 09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1292  
; TYPE: DNA  
; ORGANISM: Human  
; US-10-651-510-1

Query Match 14.8%; Score 454.6; DB 7; Length 1292;  
Best Local Similarity 62.6%; Pred. No. 7.9e-125;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

Qy 1300 TCAGAGAGTGGAGGACCTGAAGGAGAATTCGCGAAGCGACGCGCCGACAGGCGGAGA 1359  
Db 22 TCAGAGAGTGGAGAGAAATGAAGAACAGCGAGAGAGAGAAAGGCCAGAACTCTGAAA 81  
Qy 1360 TCAGAGGAGAGAGTGGCGCTGATGAACAGGATCCGGGCAATCCAACTGGGAGACG 1419  
Db 82 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTATTTCCAAA CTGGGAATTTG 135  
Qy 1420 CGCAAAATGATACGCGAAATATCAGACGACGCTGGAATTTGTGCGCTGCTCGATGGCCAG 1479  
Db 136 CCGAATGATTAAGAAATTTGGGGTACTTTGGAAATGTCATCCACTTATGACTGATC 195  
Qy 1480 CCGTCCGATGACCATCAGATCAGTGTGCGCGAAGCGTCCCATTTAGCCGCAAGGAGG 1539  
Db 196 CTATCGAAGAGACAGAAATATGTGCTGTGTTAGGAAACGCCCACTGAATAAGCAAGAT 255  
Qy 1540 TCATCGCAGAGGAGATCGATGTCATTTTCGTCGCCGCGCAGGACATGCTCATCGTGACG 1599  
Db 256 TGGCCAAAGAAAGAAATGATGTGATTTCCATTTCTAGCAAGTGTCTCCTCTTTGGTACATG 315  
Qy 1600 AGCCGCGCAGCAAGGTCCACCTCACCAAGTTCCTGAGAAACCAAAAGTTTCGCTTCGACT 1659  
Db 316 AACCCAAAGTTGAAAGTGAGCTTAACAAAGTATCTGAGAAACCAAGCATTCGCTTTGACT 375  
Qy 1660 AGCCCTTCAACGACACGTCGCAATGCCATGGTATACAAATACACAGCCAAAGCCGTTGG 1719  
Db 376 TTGCATTTGATGAACACAGCTTCGAATGAATGTTGTCTACAGGTTTCACAGCAAGGCCACTG 435  
Qy 1720 TGAAAACCATTTTCGAGGCGGAGATGGCGAGTGTGCTTCGCCCTACGCGCAGACGGATCGG 1779  
Db 436 TACAGACAACTCTTTGAAGGTGAAAGCAACTTGTGTTTGCATATGGCCAGACAGGAAGTG 495

```

QY 1780 GCAAAACGCACACCATGGCGGTGAGTTTAATGGAAGGTGCAGACTGCAGAAGCGCA 1839
DB 496 GCAAGACACATCTATGGCGGAGACCTCTCTGGAAAGCCAGAAATGCATCCAAAGGGA 555
QY 1840 TCTACGCCATGGCGCCCAAGGATGTCTTTGTGACCTCGAATATGCGCGTTACCGCGCCA 1899
DB 556 TCTATGCCATGSCCTCCCGGAGCTCTTCTCTGAAGATCAACCTCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTCTCGGCGAGTTCTTTGAGATTTCAGTGGCAAGGTCTTCGATCTTC 1959
DB 616 TGGGCTCGAAGTCTATGTGACATCTTCGAGATCTACAATGGGAAGCTGTTTGACCTGC 675
QY 1960 TGTTCGCAAGCAGCAAAATGCGCGCTCTGGAGATGTTAAACAGCAAGTGCAGGTGG 2019
DB 676 TCACAGAGAGCCAAAGCTGCGCGTCTGGAGATGGCAAGCAAGGTGCAAGTGGTGG 735
QY 2020 GACTCACCGAGAAGTGTGCGTCTGAGTGGCGTGGAGAGTACTGAAGCTCATCCAGCACGGCA 2079
DB 736 GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGGCA 795
QY 2080 ATGCTGCCGAACATCCGGCCAGAGCTGCGGCCAACTCCAAATTCGTGCGGTTCGCAAGCGG 2139
DB 796 GCGCCTGCAGAACCTCTGGGCAGACATTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT 855
QY 2140 TTTTCCAGATTGTGCTGCGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCGTCTCA 2199
DB 856 GCTTCCAAATTAATCTTCAAGTAAAG-----GAGAAATGATGGCAAGTTCCTTTGG 909
QY 2200 TCGATCTGGCGGCAATGAGCGGGCGTGGACACTCTCTCGGCCCATCGGCAGACGCGTA 2259
DB 910 TAGATCTGGCAGGAATGAGCGGCGGACACTTCCAGTGTCTGACCGGCAGACCCGCA 969
QY 2260 TGGAGGGTCCGAGATTAACTAAATCGCTGCTGGGCCCTCAAGAGTGCATTCGTGCGTTGG 2319
DB 970 TGGAGGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGGGCCCTGG 1029
QY 2320 GCARACAGTGGCCACACTTGGCTTCCGTCTCCAAACTCACCAGTGTGCGCGACT 2379
DB 1030 GACAGAACAGGCTCACACCCGTTCCGTGAGAGCAAGTGCATCAGAGTGTGAGGGACT 1089
QY 2380 COTTCATTTGGGAGAGCAAGACGTGCATGATAGCCATGATCTCGCGCGGACTTAGCT 2439
DB 1090 CTTTCATTTGGGAGAACTCTAGGACTTGCATGATGGCCACGATCTCACAGGCATAGCT 1149
QY 2440 CTTGAGACACACGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGAGCTG 2492
DB 1150 CTTGTGAATATACTTTAAACACCCCTGAGATATGCAGACAGGGTCAAGGAGCTG 1202

```

RESULT 9

```

US-10-651-510-7
; Sequence 7, Application US/10651510
; Publication No. US20040096949A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/651,510
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7

```

```

; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Human
US-10-651-510-7

```

```

Query Match 14.8%; Score 454.6; DB 7; Length 1421;
Best Local Similarity 62.6%; Pred. No. 8.3e-125;
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

```

```

QY 1300 TGAAGAGGTGGAGCGCATGGAAGGAGATCGCGAAGACGACGCGCCCGACAGGCCGAGA 1359
DB 22 TGAAGGAAGTGGAAAAAATGAAGAAACAAGCGAGAGAGAGAGGCCCGCAGAACTCTGAAA 81
QY 1360 TGAAGGAGGAGAGAGTGGCGCTGATGAACACAGGATCCGGGCAATCCAAAATCGGAGACGG 1419
DB 82 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTGTTCCTTCCAAAATCGGAAATTTG 135
QY 1420 CGCAAAATGATACGCGCAATATCAGAGACGCTGGAAATTTGTGCGCTGCTCGATGCCCAGG 1479
DB 136 CCCGAATGATTAAGAAATTTGGGCTACTTTTGGAAATGTCATCCACTTACTATGACTGATC 195
QY 1480 CCGTTCGATGACCAATCAGATCAGATGTCGCTGCGCAAGCGTCCATTTAGCGCGCAAGGAGG 1539
DB 196 CTATCGAAGAGACACAGAAATATGTGTGTGTTAGAAAACGCCCACTGAATAAGCAAGAT 255
QY 1540 TCAATCGCAAGGAGATCGATGTCTATTTTCGGTGGCGCGCAGGACATGCTCATGTGACAG 1599
DB 256 TGGCCAAAGAAAGAAATTTGATGTGATTTCCATTTCTAGCAAGTGTCTCTCTTGGTACATG 315
QY 1600 AGCGCGCAGCAAGGTTCGACCTCACCAAGTTTCTGGAGAACCAACAAGTTTCGCTTCGACT 1659
DB 316 AACCAAGTTGAAAGTGCAGCTTAACAAAGTATCTGGAGAAACCAAGCAATTCGCTTTGACT 375
QY 1660 ACGCCTTCAACGACACGTGCGCAATGCCATGCTGATATCAAAATACACAGCCAAAGCGTTGG 1719
DB 376 TTGCATTTGATGAAAACAGCTTCGAATGCAAGTTGTCTACAGTTTCCACAGCAAGGCCACTGG 435
QY 1720 TGAACAACATTTTCGAGGGCGGAAATGGCGAGGTGCTTCGCGCTACGGCGCAGACGGGATCGG 1779
DB 436 TACAGACAATCTTTTGAAGGTGGAAAAGCAACTTTGTTTTCATATATGGCCAGACAGGAAGTG 495
QY 1780 GCAAAACGCACACCATGGCGGTGAGTTTAAATGAAAAGGTGCAGAGACTGCAAGAAACGGCA 1839
DB 496 GCAAGACACATCTATGCGGCGGAGACCTCTCTGGGAAAGCCAGAGATGCATCCAAAGGGA 555
QY 1840 TCTACGCCATGGCGGCCAAGGATGTCTTTGTGACCCCTGAAATATGCGCGGTTACCGCGCCA 1899
DB 556 TCTATGCCATGGCGCTCCCGGAGCGTCTTCTCTGMAAGAAATCAACCTGCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTCTTTGAGATTTTACGTGGCAAGGTCTTCGATCTTC 1959
DB 616 TGGGCCCTGGAAGTCTATGTGACATTTCTTCAGATCTTACAATGGGAAGCTGTTTGACCTGC 675
QY 1960 TGTCCGACAAGCAGAAACTGCGGCTCTCGAGGATGGTAAACAGCAAGTGCAGGTGTGG 2019
DB 676 TCAACAAGAGGCCAAGCTGCGGTGCTGGAGGATGCAAGCAACAGGTGCAGAGTGTGG 735
QY 2020 GACTCACCGAGAAGTGTGCGTGGCGTTCGAGGAGTACTGAAAGTCTCATCCAGCACGGCA 2079
DB 736 GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGGCA 795
QY 2080 ATGCTGCCCAACATCCGGCCAGACGTCGGGCCAACTCCAAATTCGTGCGGTTCGACACCGG 2139
DB 796 GCGCCTGCAGAACCTCTGGGCAGACATTTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT 855
QY 2140 TTTTCCAGATTGTGCTGCGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCCTCGTTCA 2199
DB 856 GCTTCCAAATTAATTCCTCGAGCTTAAGG-----GAGAAATGCAATGGCAAGTTCCTTTGG 909
QY 2200 TCGATCTGGCGGCAATGAGCGGGCGTGGACACTTCTCTCGGCGGATCGGCAGACGCGTA 2259
DB 910 TAGATCTGGCAGGAAATGAGCGGCGGCGACACTTCCAGTGTGCTGACCGGCAGACCCGCA 969

```



QY 2260 TGGAGGGTCCGAGATTAAACAATCGCTGTGGCCCTCAGGAGTGCAATTCGTGCGTTGG 2319  
Db 970 TGGAGGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTCGAAGAGTGATCATCAGGGCCCTGG 1029  
QY 2320 GCAAAACAGTCCGCCACTTCGCCCTTCGGTCTCTCCAACTCCACCCAGGTGCTCGCGGACT 2379  
Db 1030 GACAGAACAGGCTCACACCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGGACT 1089  
QY 2380 CGTTCAATTGCGGAGAGAGCAAGACGTGCGATGATAGCATGATCTCCCGGGGACTTAGCT 2439  
Db 1090 CCTTCAATTGCGGAGAACTCTAGGACTTGCATGATTCACGATCTCACAGGCGATAAGCT 1149  
QY 2440 CCTGCGAGCACGCTCAACACGCTCGCTATCGGATCGTGTCAAGGAGCTG 2492  
Db 1150 CCTGTGAATATACTTTAAACACCCCTGAGATATGCACAGGGGTCAAGGAGCTG 1202

## RESULT 10

US-10-651-510-13  
; Sequence 13, Application US/10651510  
; Publication No. US20040096949A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042A  
; CURRENT APPLICATION NUMBER: US/10/651,510  
; CURRENT FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US/09/724,215  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; PRIOR APPLICATION NUMBER: 09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 2172  
; TYPE: DNA  
; ORGANISM: Human  
US-10-651-510-13

Query Match 14.8%; Score 454.6; DB 7; Length 2172;  
Best Local Similarity 62.6%; Pred. No. 1.1e-124;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
QY 1300 TGAAGAGTGGAGCGACTGAAGGAGAAATCGCGAGAGCGACGCGCCGACAGGCGGAGA 1359  
Db 578 TGAAGGAAAGTGGAAAAAATCAAGAAACAAGCGAGAGAGAAAGGCGCCAGAACTCTGAAA 637  
QY 1360 TGAAGGAGGAGAGTGGCGCTGATCAACAGGATCCGGGCAATCCAACTGGGAGACGG 1419  
Db 638 TGAGAAATGAAGAG-----CTCAGAGATGACAGTAGTGTTCAAAATGGGAATTG 691  
QY 1420 CGCAATGATACCGCAATATCAGAGCACGCTGGAATTTGTGCGCTGCTCGATGGCCAGG 1479  
Db 692 CCGGAATGATTAAGAAATTTGGGCTACTTTGGAATGTATCCACTTACTATGACTGATC 751  
QY 1480 CGGTGATGACCATCATGATCAGATGTCGTGCGCAAGCGTCCCAATAGCCGCAAGGAGG 1539  
Db 752 CTATCGAAGAGCACAGATATGTGTCTGTGTAGGAAAGCCCACTGAATAAGCAAGAAT 811  
QY 1540 TCAATCGCAGGAGATCGATGTCATTTCCGTCCGCGGCAAGGACATGCTCATCGTCAGC 1599  
Db 812 TGCCCAAGAAAGAAATTTGATGTAATTCATTCTTAGCAAGTGTCTCTCTTGGTACATG 871  
QY 1600 AGCCGGCGCAGGAGTCCGACCTCACCAGTTCCTCGGAGAACCAAGATTTGCTTTCGACT 1659  
Db 872 AACCCAAAGTTGAAGTGGACTTAAACAAAGATATCTGGAGAACCAAGCATTTCTGCTTGA 931

QY 1660 AGCCCTTCAACGACACGTCGGACATGCCATGGTATACAAAATACACAGCCCAAGCCCTGG 1719  
Db 932 TTGCATTTGATGAAAACAGCTTCGAATGAAGTTGTCTACAGGTTTACAGCAAGCCACTGG 991  
QY 1720 TGAATAACCATTTTCGAGGGCGGAATGCGCAGCGTCTTCGCTTACGGCCAGAGCGGATCG 1779  
Db 992 TACAGCAATCTTTGAAGGTGGAAAGCAACTTTGTTTGCATATGCCCAGACAGGAAGTG 1051  
QY 1780 GCAAAAACGACACCAATGGGCGGTGAGTTTAAATGGAAGGTGCGAGGACTGCAAGAACGGCA 1839  
Db 1052 GCAAGACACATACTATGTGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCAATCCAAAGGA 1111  
QY 1840 TCTACGCCATGCGGGCCAAAGGATGCTTTGTGACCTGTAATATATGCGCGTTACCGGCCCA 1899  
Db 1112 TCTATGCCATGGCCTCCCGGACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1171  
QY 1900 TGAATCTAGTCTCTCGGCGAGTTTCTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC 1959  
Db 1172 TGGGCTTGAAGTCTATGTGACATCTTTCGAGATCTACATGGGAAGCTGTTTGACCTGC 1231  
QY 1960 TGTCCGACAAAGCAGAAATTCGCGCTCTGGAGAGTGGTAAACAGCAAGTGCAGGTGGTG 2019  
Db 1232 TCAACAAAGAAAGCCAAAGCTGCGCTGCGAGGATGGCAAGCAACAGGTGCAAGTGGTG 1291  
QY 2020 GACTCACCGAGAAAGTGGTGGATGGGTCGAGGAGTACTGAAGCTCATCCAGCAGCGCA 2079  
Db 1292 GGTCTGAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGCA 1351  
QY 2080 ATGTGCCCCGAACATCCGSCCAGAGCTGGGCCAACTCCAAATTCGTCGCTGCGGCTGCACGCG 2139  
Db 1352 GCGCTGACAGAACCTCTGGGCGAGACATTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT 1411  
QY 2140 TTTTCCAGATTTGCTGCGCGCGCGAGGCTCGACGAAAGATCCATGGCAAGTTCCTCGTTCA 2199  
Db 1412 GCTTCCAAATTAATCTTCGAGCTAAAG-----GAGATGATGCAAGTTCCTCTTGG 1465  
QY 2200 TCGATCTGCGGGCAATGAGCGGGCGTGGAACACTTCTCGCGCGATTCGGCAGACGCGTA 2259  
Db 1466 TAGATCTGGCAGGAAATGAGCGGAGCGCGACACTTCCAGTGTGACCCGCGAGACCGCA 1525  
QY 2260 TGGAGGTGCGGAGATTAACAATCGCTGCGCCCTCAAGAGGTGCAATTCGTCGCTGG 2319  
Db 1526 TGGAGGCGCGAGAAATCAACAAGAGTCTCTTAGCCCTGAGGAGTGATCAGGGCCCTGG 1585  
QY 2320 GCAAAACAGTCCGCCCACTTTCGCTTCCGCTTCTCCAACTCACCAGGTGCTGCGGACT 2379  
Db 1586 GACAGAACAAAGCTCACACCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1645  
QY 2380 CTTTCAATGGCGAGAGAGCAAGAGCTGATGATAGCCATGATCTCGCCGCGGACTTAGCT 2439  
Db 1646 CTTTCAATGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCAGGCGATAAGCT 1705  
QY 2440 CTTGCGAGCACAGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492  
Db 1706 CCTGTGAATATACTTTAAACACCCCTGAGATATGCACAGGGTCAAGGAGCTG 1758

## RESULT 11

US-10-240-965-208  
; Sequence 208, Application US/10240965  
; Publication No. US20030165924A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: SHIFFMAN, Dov  
; APPLICANT: SOMOGYI, Roland  
; APPLICANT: LAWN, Richard M.  
; APPLICANT: SEILHAMER, Jeffrey J.  
; APPLICANT: PORTER, Gordon J.  
; APPLICANT: MIKITA, Thomas  
; APPLICANT: TAI, Julie  
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION  
; FILE REFERENCE: PA-0025 PCT  
; CURRENT APPLICATION NUMBER: US/10/240,965

```
/ CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 208
; LENGTH: 2865
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 331025.1
US-10-240-965-208

Query Match
Best Local Similarity 14.7%; Score 453; DB 6; Length 2865;
Matches 746; Conservative 0; Mismatches 435; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGGAGTGAAGGAGATCGCGAGAAGCGCGCGCCGACAGAGCGCGAGA 1359
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
675 TGAAGGAGTGGAAAAATGAAGAACAGAGGAGAGAGAGAGGCGCCAGRACTCTGAAA 734
QY 1360 TGAAGGAGAGAGGTGGCGCTGATGAACACAGAGTCCGGGCGAATCAAACTGGGAGACGG 1419
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
735 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTATTTTCCAAACTGGGAATTTG 788
QY 1420 CCAGAAATGATACGCAATATACAGACAGCGCTGGAATTTGTCCGCTGCTCGATGCCAGG 1479
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
789 CCGGAATGATTAAGAAATTCGGGTACTTTGGAATGTCATCCACTTACTATGACTGATC 848
QY 1480 CGCTCGATGACCATCAGATCAGTGTGCGTGCAGCAAGCGTCCCATTTAGCCGCAAGGAGG 1539
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
849 CTATCGAAGAGCACAGATATGTGCTGTGTTAGGAACGCCACTGAATGAAGCAAGAT 908
QY 1540 TCAATCGAAGAGATCGATGCTATTTTCGGTGCAGCAAGACATGCTCATCGTGCACG 1599
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
909 TGGCAAGAAAGAAATGATGATGATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 968
QY 1600 AGCCGCGCAGCAAGTTCGACCTCACCAAGTTCTCGAGAACACCAAGTTTCGCTCGACT 1659
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
969 AACCCAAAGTTGAAGTGACTTAACAAAGATCTGGAGAACCAAGCATTCGCTTTGACT 1028
QY 1660 AGCCTTCAACGACACGTGCGCAATGCCATGGTATACAAATACACAGCCCAAGCGTTGG 1719
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1029 TTGCATTTGATGAACACAGCTTCGAATGAAGTTGTCTACAGTTTCACAGCAAGGCCACTGG 1088
QY 1720 TGAAGAACATTTTCAGGCGGGAATGCGGACGTGCTTCGCTACGCGCAGACGGGATCGG 1779
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1089 TACAGACAATCTTTGAAGGTGGAAAGCAACTTGTGTCATATGCGCCAGACAGGAAGTG 1148
QY 1780 GCAAAACGCACACCATGGCGGTGATTTAATGGAAGGTGCAGGACTGCAAGAACGCA 1839
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1149 GCAAGACACATCTATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGGA 1208
QY 1840 TCTAGCCATGGCGGCAAGGATGCTTTGTGACCTCAATATGCGCGTTTACCGCGCA 1899
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1209 TCTATGCCATGGCTCCCGGAGCTCTCTCTCTGAAGAAATCAACCTGCTACCGGAAGT 1268
QY 1900 TGAATCTAGTCTGCTCGGCGAGTTCTTTTGAATTTACAGTGGCAAGGTCTTCGATCTTC 1959
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1259 TGGGCTTGAAGTCTATGTGACATCTCTCGAGATCTACAATGGGAAGGTGTTGACCTGC 1328
QY 1960 TGTCCGCAACAGCAAGAACTGCGCGTCTGGAGGATGTTAAACAGCAAGTGCAGTGTGG 2019
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1329 TCAACAGAGGCGAGCTGCGGTCTGGAGGAGCGGCAACAGGTGCAAGTGGTGG 1388
QY 2020 GACTCACCGAAGGTGGTTCGATGCGGTGCGAGGAGTACTGAAGTCTATCCAGCACGCGCA 2079
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1389 GGCTGCAGGACATCTGCTTAACTCTGCTGATGATGATCATCAAGATGATGCACATGGCA 1448
QY 2080 ATGTCGCCGGAACATCCGCGCAGAGCTCGGCGCACTCAATTCGTGCGGTTTCGACGCGG 2139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1449 GCGCTGCGAGAACTCTGGGAGAGCATTTGCGCAACTCCAAATTCCTCCGCTCCACGCGT 1508
```

```
2140 TTTTCCAGATTGTCTGCGCGCGCAGGCTCGACGAAGATCCATGCGAAGTTCTCGTTCA 2199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1509 GCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGAAATGCGGCAAGTTCTTTGG 1562
QY 2200 TCGATCTGGCGGCAATGAGCGGGCGGTGGACACTTCTCTCGGCGGATCGGCGAGCGGTA 2259
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1563 TAGATCTGGCAGGGAATGAGCGGCGGAGACACTTCCAGTGTCTGACCGGAGACCGCA 1622
QY 2260 TGGAGGTTGCCGAGATTAAACAAATCGTGTGGGCCCTCAAGAGTGCATTCGTGCTGG 2319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1623 TCGAGGCGCGAGAAATCAACAAAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGCGCCCTGG 1682
QY 2320 GCAACAGTCTGGCCCACTTTCGCTTCCGCTCTCCAACTCACCAGGTGCTGCGGACT 2379
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1683 GACAGAAACAGGCTCACACCCCTTCGCTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1742
QY 2380 CGTTCAATGGCGAGAGCAAGACGTCATGATAGCATGATCTCGCGGAGACTTAGCT 2439
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1743 CTTTCAATGGGAGAACTCTAGGACTTGCATGATGCGACGATCTCACCAGGCAATAGCT 1802
QY 2440 CTTGCGAGCACACGCTCAACACGCTGCGCTATGCGGATTCGTTCAAGAGCTG 2492
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1803 CCTGTGAATATATCTTTAAACACCCCTGAGATATGAGACAGGCTCAAGGCTG 1855

RESULT 12
US-10-510-3
; Sequence 3, Application US/10651510
; Publication No. US20040096949A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/651,510
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1175
; TYPE: DNA
; ORGANISM: Human
US-10-651-510-3

Query Match
Best Local Similarity 14.6%; Score 449.8; DB 7; Length 1175;
Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

QY 1402 ATCCAACTGGGAGACGGCGCAATGATACGCAATATCAGAGCACGCTGGAATTTGTGC 1461
Db 1 ATCCAACTGGGAATTTGCGCGAATGATTAAAGAAATTCGGGCTACTTTGGAATGTCAATC 60
QY 1462 CGTGTCTGATGGCAGGCGCTCGATGATCAGATCAGATGTCGCTGCGGAGGCTC 1521
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 CACTTACTATGATGATCTCTATCGAAGAGCACAGAAATATGTGCTGTGTGTTAGGAAACGCC 120
QY 1522 CCATTAGCCGCAAGGAGGTCAATCGCAAGGAGATCGATGTCATTTCCGTCGCGCGCAAGG 1581
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 CACTGAATAGCAAGAAATTTGGCCCAAGAAATGATGATGATTTCAATTCCTAGCAAGT 180
QY 1582 ACATGCTCATCTGTCAGAGCGCGGAGCAAGGTTCGACCTCAACCAAGTTCTGAGAACCC 1641
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 GTCTCTCTTGGTACATGAACCCCAAGTTGAAAGTGGAGCTTAAACAAAGTATCTGAGAACCC 240
```

```
QY 1642 ACAAGTTTCGTTCCGACTACGCTTCAACGACACGTCGCGCAATGCGCATGGTATACAAAT 1701
Db 241 AAGCATTTCTGTTGACTTTGCAATTTGATGAACACAGCTTCCGAATGAAGTTGTCTACAGGT 300
QY 1702 ACACAGCCAAAGCCGTTGGTGAACCAATTTTCAGAGGCGGAATGGCGACGTCTCGCCT 1761
Db 301 TCACAGCAAGGCCACTGCTACAGCAATCTTTGAGGTGGAAGCACTGTTGTTGTCAT 360
QY 1762 ACGCCAGACGGGATCGGGCAAAACGACACCAATGGCGGTGAGTTTAAATGGAAGGTGC 1821
Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATATATGGCGGAGACCTCTCTGGGAAGCCC 420
QY 1822 AGCACTGCAAGACGGCATCTAGCCCATGGCGGCAAGATGCTTTGTGACCCCTGAATA 1881
Db 421 AGAATGCATCCAAAGGGATCTATGCCATGGCCTCCCGGAGCGTCTTCCTCTGAAGAATC 480
QY 1882 TGCGCGGTTACCGCGCCATGAATCTAGTCTGCTCGGCCAGTTCTTTGAGATTTACAGTG 1941
Db 481 AACCTGCTACCGAAGTTGGGCTTGAAGTCTATGTGACATTTCTCGAGATCTCAATG 540
QY 1942 GCAAGTCTTTCGATCTTCTGTCGCAAGCAGAAACTGCGCGTCTCTGAGGATGGTAAAC 2001
Db 541 GGAAGCTGTTGACCTGCTCAACAGAGGCCAAGCTGCGGTGCTGGAGATGGCAAGC 600
QY 2002 AGCAATGCAAGTGGTGGGACTCACCGAAGGTGTCAGATGGCGGTGCGAGGATGCTGA 2061
Db 601 AACAGTGCAGAGTGGTGGGCTGTCAGAGGATCTGGTTAACTCTGCTGATGATGTCATCA 660
QY 2062 AGCTCATCCAGCACGCAATGCTGCCGCAACATCCGCGCAGAGCTGCGGCCCACTCAAT 2121
Db 661 AGATGATGACATGGGCGGCGCTGAGAACCTCTGGGAGACATTTGGCCAACTCCAAT 720
QY 2122 CGTGGGTTTCGACGCGCTTTTCCAGATTTGCTGGCGGCGCAGGCGCTCGACGAAGATCC 2181
Db 721 CCTCCGCTCCCAAGGCTGCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGATGC 774
QY 2182 ATGGCAAGTTCTGTTTCATGATCTGGCGGCAATGAGCGGCGGTGGACACTTCTCTCGG 2241
Db 775 ATGGCAAGTTCTGTTTGGTAGATCTGGCAGGGAATGAGCGGCGGCACTTCCAGTG 834
QY 2242 CCGATCGGACAGCGCTATGAGGGTGGCGAGATTAAACAAATCGCTGCTGGCCCTCAAG 2301
Db 835 CTGACGGGACAGCCCGATGAGGGGCGAGAAATCAACAGAGTCTCTTAGCCCTGAAG 894
QY 2302 AGTGCAATTCGTGGTGGGCAACAGTCGCGCCCACTTGGCCCTTCGCTGCTCAAACTCA 2361
Db 895 AGTGCAATCAGGCGCTGGGACAGAACAGGCTCACACCCGTTCCGTGAGAGCACTGA 954
QY 2362 CCGAGTGTGCGGACTCGTTCAATGGGAGAGCAAGACGTGCATGATAGCCATGA 2421
Db 955 CACAGTGTGAGGGAATCTTCAATGGGAGAACTCTAGGACTTGCATGATGGCCACGA 1014
QY 2422 TCTCGCGGACTTACCTCTGGGACACAGCTCAACAGCTGCGCTATGCGATGCGTG 2481
Db 1015 TCTCACAGGCATAAGCTCTCTGTAATATACTTTAAACCCCTGAGATATGACAGAGG 1074
QY 2482 TCAAGGAGCTG 2492
Db 1075 TCAAGGAGCTG 1085
```

## RESULT 13

```
US-10-651-510-9
; Sequence 9, Application US/10651510
; Publication No. US2004009694A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/651,510
```

```
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Human
US-10-651-510-9
```

```
Query Match 14.6%; Score 449.8; DB 7; Length 1304;
Best Local Similarity 64.0%; Pred. No. 2.2e-123;
Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;
QY 1402 ATCCAACTGGGAGACGGCGCAAAATGATACGCGAATATCAGACACGCTGGAATTTGTGC 1461
Db 1 ATGCAAACTGGGAATTTGCCGAATGATTAAGAATTTTCGGCTACTTTTGGAAATGTCATC 60
QY 1462 CGCTGCTCGATGGCCAGGCGCTCGATGACCATCAGATCAGATGTCGTGCGTGGCAAGCGTC 1521
Db 61 CACTTACTATGACTGATCTCTATCGAAGAGCACAGATATATGTCTGTGTAGGAAGGCC 120
QY 1522 CCATTAGCCGCAAGGAGTCAATCGCAAGAGATCGATGTCAATTCGGTGGCGCGCAAGG 1581
Db 121 CACTGAATAAGCAAGAAATTTGGCCAAAGAAATTTGATGTGATTTCCATTTCTTAGCAAGT 180
QY 1582 ACATGCTCATCTGTGACGAGCGCGCAGCAAGCTCGACCTCACCAAGTTCTTCGAGAAC 1641
Db 181 GTCTCTCTTGTGTACATGAAACCAAGTTGAAAGTGAGCTTAACAAAGTATCTCGAGAAC 240
QY 1642 ACAAGTTTCGCTTCGACTACGCTTCAACGACACGTCGCACAATGCGCATGGTATACAAAT 1701
Db 241 AAGCATTTCTGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 300
QY 1702 ACACAGCCAAAGCGTTGGTGAACCAATTTTCAGGGCGGAATGGCGACGTCTCGCCT 1761
Db 301 TCACAGCAAGGCCACTGCTGACAGCAATCTTTGAAAGTGGAAAAAGCACTTGTGTTTGCAT 360
QY 1762 ACGCCAGACGGGATCGGGCAAAACGACACCACTGGCGGTGAGTTTAAATGGAAGGTGC 1821
Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATACTATGGCGGAGACCTCTCTCGGAAAGCCC 420
QY 1822 AGCACTGCAAGACGGCATCTACGCCATGGCGGCAAGGATGCTTTTGTGACCCCTGAATA 1881
Db 421 AGAATGCATCCAAAGGGATCTATGCCATGGCCTCCCGGAGCGTCTTCCTCTCGAAGATC 480
QY 1882 TGCGCGGTTACCGCGCCATGAATCTAGTCTCGTCTCGGCCAGTCTTTTGTGAGATTTACAGTG 1941
Db 481 AACCCCTGCTACCGGAAGTTGGGCTTGAAGTCTATGTGACATTTCTCGAGATCTCAATG 540
QY 1942 GCAAGTCTTTCGATCTTCTGTCGCAAGCAGAAACTGCGCGTCTCTGAGGATGGTAAAC 2001
Db 541 GGAAGCTGTTGACCTGCTCAACAGAGGCCAAGCTGCGGTGCTGGAGATGGCAAGC 600
QY 2002 AGCAAGTGCAGGTGGTGGGACTCACCGAAGGTGCTCGATGGCGGTGCGAGGATGCTGA 2061
Db 601 AACAGTGCAGAGTGGTGGGCTGTCAGAGGATCTGGTTAACTCTGCTGATGATGTCATCA 660
QY 2062 AGCTCATCCAGCACGCAATGCTGCCGCAACATCCGCGCAGAGCTGCGGCCCACTCAAT 2121
Db 661 AGATGATGACATGGGCGGCGCTGAGAACCTCTGGGAGACATTTGGCCAACTCCAAT 720
QY 2122 CGTGGGTTTCGACGCGCTTTTCCAGATTTGCTGGCGGCGCAGGCGCTCGACGAAGATCC 2181
Db 721 CCTCCGCTCCCAAGGCTGCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGATGC 774
```

```
QY 2182 ATGGCAAGTTCCTGTTTCATCGATCTGGCGGGCAATGAGCGGGCGTGGACACTTCTCTCGG 2241
Db      |||||
775 ATGGCAAGTTCCTTTTGGTAGATCTGGCAGGGAATGAGCGGGCGGACACTTCCAGTG 834
QY 2242 CCGATCGGCAGACGGGTATGAGGGTGGCGAGATTAACTAAATCGCTGTGGCCCTCAAGG 2301
Db      |||||
835 CTGACCGGCAGACCCGCACTGGAGGGCGCAGAAATCAACAAGAGTCTCTTTAGCCCTGAAG 894
QY 2302 AGTGCAATTCGCTGTTGGCAACAGTGGGCCCACTTGCCCTTCCTCGTGTCTCCAACTCA 2361
Db      |||||
895 AGTGCAATCAGGCCCTGGGACAGAACAGGCTCACACCCGTTCCGTGAGAGCAAGCTGA 954
QY 2362 CCCAGTGTGTCGCGACTCGTTCAATGGCGAGAGCAAGACGTGTCATGATAGCCATGA 2421
Db      |||||
955 CACAGTGTGAGGGACTCTTCAATTTGGGAGAACTCTAGGACTTGCATGATTGCCAGA 1014
QY 2422 TCTCCCGGGACTTACGTCTCTCGGAGCACACGCTCAACGCTGCGGCTATGCGGATCGTG 2481
Db      |||||
1015 TCTCACCAGGCATAGCTCTCTGTAATATACCTTTAAACACCCCTGAGATATGCAGACAGG 1074
QY 2482 TCAAGGAGCTG 2492
Db      |||||
1075 TCAGGAGCTG 1085

RESULT 14
US-09-954-456-77
; Sequence 77, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 77
; LENGTH: 2740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-77

Query Match 14.6%; Score 449.8; DB 3; Length 2740;
Best Local Similarity 62.4%; Pred. No. 3.4e-123;
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGGATCCGAGAGCGACGCGCCCGCAGCGCCGAGA 1359
Db      |||||
638 TGAAGAGGTGGAAAAAATGAAAGAACAGCGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGG 697
QY 1360 TGAAGGAGGAGAGGTGGCGCTGTATGAACAGGATCCGGGCAATCCAACTGGGAGACGG 1419
```

```
Db      |||||
698 TGAGATGAAGAGAG-----CTCAGGAGTATGACAGTAGTTTTTCCAACTGGGAAATTG 751
QY 1420 CGCAAAATGATACGGCAATATCAGAGCAGCGTGGAAATTTGTCCCGTGTCTCAATGGCCAGG 1479
Db      |||||
752 CCCGAATGATTAAGAATTTCCGGGCTACTTTGGAAATGTCAATCCACTTACTATGACTGATC 811
QY 1480 CCGTCGATGACCATCAGATCAGATGTGCGTGGCAAGCGTCCCATAGCCGCAAGGAGG 1539
Db      |||||
812 CTATCGAAGAGCAGAGAAATATGTGTCTGTGTAGGAAACGCCCACTGAATTAAGCAAGAAT 871
QY 1540 TCAATTCGCAAGGAGATCGATGTCAATTTTCGGTGGCGCGCAAGGACATGCTCATCGTGACG 1599
Db      |||||
872 TGGCCAGAGAAATTTGATGTGATTTCCATTCCTAGCAAGTGTCTCTCTTGTGTACATG 931
QY 1600 AGCCCGCAGCAAGGTGCACTCAACCAAGTTCCTGGAGAACCAACAAGTTTCGTTTCGACT 1659
Db      |||||
932 AACCCAAAGTTTGAAGTGGACTTAAACAAAGTATCTGGAGAACCAAGCAATTCCTGTTGACT 991
QY 1660 ACGCCTTCAACGACACGTCGACAAATGCCATGGGTATACAAATACACAGCCAAAGCCGTTGG 1719
Db      |||||
992 TTGCAATTTGATGAACAGCTTCGAATGAAGTTGTCTACAGGTTCAACAAAGGCCACTGG 1051
QY 1720 TGAATAACCATTTTCAGGGCGGAAATGGCGAGCTTTCGCTTACGGCCAGACGGGATCGG 1779
Db      |||||
1052 TACAGACAACTCTTTGAAGGTGGAAGCAACTTGTTCATATATGGCCAGACAGGAAGTG 1111
QY 1780 GCAAAAACGACACCAATGGCGGTGAGTTTAATGAAGAGGTGCAGGACTGCAAGAACGGCA 1839
Db      |||||
1112 GCAAGACACATACTATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGGA 1171
QY 1840 TCTAGCCCATGGCGGCAAGGATGCTTTGTGACCTTGAAATATCCGCGTTTACCGCGCCA 1899
Db      |||||
1172 TCTATGCCATGGCCCTCCCGGAGCGTCTCTCTTGAAGAATCAACCTCTGCTACCGGAAGT 1231
QY 1900 TGAATCTAGTGTCTCGGCCAGATTTCTTTGAGATTTTACAGTGGAAGGTCTTTGATCTTC 1959
Db      |||||
1232 TGGSCCTTGAAGTCTATGTGACATTTCTCGAGATCTACAAATGGGAAGCTGTTGACCTGC 1291
QY 1960 TGTCCGACAAGCAGAAATCTGCGGTCTCGAGGATGTAAACAGCAAGTGCAGGTGTGTGG 2019
Db      |||||
1292 TCAACAAGAAAGGCCAAGCTGCGGTGCTGGAGGACGGCAAGCAACAGGTGCAAGTGTGG 1351
QY 2020 GACTCACCGAGAGGTGCTCGATGGCGTCGAGGAGGTACTGAAGCTCATCCAGCACGGCA 2079
Db      |||||
1352 GGCTGCAGAGGACATCTGTTAACTCTGTGATGATCATCAAGATGCTCAATATGGGCA 1411
QY 2080 ATGCTGCCGAACATCCGGCCAGACGTGGCCCAACTCCAATTCGTCGCTTCGCACGCG 2139
Db      |||||
1412 GCGCTTCAGAAACCTCTGGGCAGACATTTGCCAACTCCAATTCCTCCCGCTCCACGCGT 1471
QY 2140 TTTTCCAGATTTGCTCGCGCGCAGGGCTCGAGCAAGATCCATGGCAAGTTCTCGTTCA 2199
Db      |||||
1472 GCTTCCAAATTTATCTTCGAGCTTAAAGS-----GAGAATGCAATGGCAAGTTCTCTTTGG 1525
QY 2200 TCGATCTGGCGGGCAATGAGCGGGCTGGACACTTCTCCGCGGATCGGAGACGGGTA 2259
Db      |||||
1526 TAGATCTGGCAGGGAATGAGCGGGCCAGACACTTCCAGTGTGACCGGCAAGACCCGCA 1585
QY 2260 TGGAGGTTGCCGAGATTAAACAAATTCGCTGCTGGCCCTCAAGGAGTGCATTCGTGCGTTGG 2319
Db      |||||
1586 TGGAGGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGGAGTGCATCAGGGSCCCTGG 1645
QY 2320 GCAAAAGTGGGCCCACTTCCCTTCGCTGTCTCCAAATCTACCCAGGTGCTGGCGACT 2379
Db      |||||
1646 GACAGAAACAAGGCTCAACACCCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1705
QY 2380 CGTTTATTTGGCAGAGAGCAGACGTGATATAGCATGATCTCCCGGGACTTAGCT 2439
Db      |||||
1706 CCTTCATTTGGGAGAACTCTAGGACTTGTGATGATTTGCCAGATCTCACAGGCATAGCT 1765
QY 2440 CTTGCGAGCACACGCTCAACACGCTGCGCTATCGGATCGGTGTCAGGAGCTG 2492
```

Db 1766 CCTGTGATATACCTTTAAACACCCCTGAGATATGACAGACAGGGTCAAGGAGCTG 1818

## RESULT 15

US-09-954-456-710  
; Sequence 710, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 710  
; LENGTH: 2740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-710

Query Match 14.6%; Score 449.8; DB 3; Length 2740;  
Best Local Similarity 62.4%; Pred. No. 3.4e-123;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;  
QY 1300 TGAACAGGTGGAGGACTGAAGGAGAAATCGGAGAGCGACGCGCCGACAGCGCGAGA 1359  
Db 638 TGAGGAGGTGGAAAATAATGAAGAACAGCGAGAGAGAGAGAGGCGCCAGAACTCTGAAA 697  
QY 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACACAGGATCCGGGCAATCCAAACTCGGAGACGG 1419  
Db 698 TGAGATGAAGAGAG-----CTCAGGATATGACAGATAGTTTCCAACTGGGAATTG 751  
QY 1420 CGCAATGATACCGGAATATCAGAGACGCTGGAATTTGTGCGCTGCTCGATGGCCAGG 1479  
Db 752 CCGGAATGATTAAGAAATTCGGGCTACTTTGGAATGTCTATCCACTTACTATGACTGATC 811  
QY 1480 CCGTCGATCAGATCAGATGCGTGGCGCAGGCTCCCATAGCCGCGAGGAGG 1539  
Db 812 CTATCGAAGAGCAGAAATATGTGTGTGTAGGAAACGCCCACTGAATAAGCAAGAAT 871  
QY 1540 TCAATCGCAGGAGATCGATGTCATTTCCGTGCGCGCAAGGACATGCTCATCGTGACG 1599  
Db 872 TGGCCAAAGAAAGAAATTTGATGTGATTTCCATTCCTAGCAAGTGTCTCCTCTTGGTACATG 931  
QY 1600 AGCCGCGCAGCAGGTGCGACTTCAACCAAGTTCTGTGAGAACCAAGTTTCTGCTTGGACT 1659  
Db 932 AACCCAAAGTTGAAGTGGACTTAAACAAAGTATCTGGAGAACCAAGCATTTCTGCTTGGACT 991  
QY 1660 AGCCCTTCAACGACAGTGGGACATGCCATGGTATACAAATACACAGCCCAAGCCGTTGG 1719  
Db 992 TTGCATTTGATGAACAGCTTCGAATGAAGTTGTCTACAGGTTTACAGCAAGGCCCACTGG 1051

QY 1720 TGAACACCACTTTTCGAGGCGGAATGCGACGCTGCTTTCGCTACGCGCCAGACGGGATCGG 1779  
Db 1052 TACAGACAATCTTTGAAGGTGGAAGCAACTTGTGTTTGCATATGCGCAGACAGGAGTG 1111  
QY 1780 GCAAAACGACACCATGCGGCGGTGAGTTTAAATGGAAGGTGAGGAGCTGCAAGAACGGCA 1839  
Db 1112 GCAAGACACATATATGCGGCGGAGACCTCTCTGGAAGGCCAGAAATGCAATCCAAAGGGA 1171  
QY 1840 TCTACGCCATGCGGCGCAAGGATGCTTTGTGACCTGACCTGAAATATGCGCGCTTACCGCGCCA 1899  
Db 1172 TCTATGCCATGCGCTCCGCGAGCTTCTCTCTGAAGAAATCAACCTGCTACCGGAAGT 1231  
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTTCTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC 1959  
Db 1232 TGGGCTGGAAGTCTATGTGACATTTCTTCGAGATCTACAATGGGAAGCTGTTTGAACCTGC 1291  
QY 1960 TGTCCGACAGCAGAACTGCGGCTCTGAGGAGTGTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 1292 TCAACAAAGAAAGCCAAAGCTGCGCGTGTGAGGAGCGCAAGCAAGGTGCAAGTGGTGG 1351  
QY 2020 GACTCACCGAGAAAGGTGGTTCGATGGCGTCGAGGAGTACTGGAAGCTCATCCAGCAGCA 2079  
Db 1352 GGTGCGAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGCTCGACATGGGCA 1411  
QY 2080 ATGCTGCCCGAACATCCGGCCAGAGCTCGGCCCAACTCCAAATTCGTCGCGTTCGACGCGG 2139  
Db 1412 GCGCTCGACAGACCTCTGCGGAGACATTTTGGCAACTCCAAATTCCTCCCGCTCCACGCGT 1471  
QY 2140 TTTTCCAGATTTGCTGCGGCGCGAGGCTCGACGAGATCCATGGCAAGTTCTCGTTCA 2199  
Db 1472 GCTTCCAAATTTATTTCTCGAGCTAAAGG-----GAGAAATGCAATGGCAAGTTCTCTTTGG 1525  
QY 2200 TCGATCTGCGCGGCAATGAGCGGGGTGCGACACTTTCCTCGGCGGATCGGACAGCGGTA 2259  
Db 1526 TAGATCTGCGAGGAAATGAGCGGCGAGACACTTTCAGTGTCTGACCGCGAGACCCGCA 1585  
QY 2260 TGGAGGCTCCGAGATTAACAAATCGCTGCTGCGCCCTCAAGGAGTGCATTTCGTGCGTTGG 2319  
Db 1586 TGGAGGCGCAGAAATCAACAAGATCTCTTAGCCCTGAAGGAGTGCATCAGGGCCCTGG 1645  
QY 2320 GCAACAGTCCGCGCCACTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2379  
Db 1646 GACAGAACAAAGCTCACACCCGTTCCGTTGAGAGCAAGCTGACACAGGTGCTGAGGGA 1705  
QY 2380 CGTTCAATTGGCGAGAGAGCAAGAGCTGCATGATAGCCATGATCTGCGCGGACTTAGCT 2439  
Db 1706 CTTTCATTTGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCAGGCAATAGCT 1765  
QY 2440 CTTGCGAGCACACGCTCAACACGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492  
Db 1766 CTTGTGATATACTTTAAACACCCCTGAGATATGACAGACAGGCTCAAGGAGCTG 1818

Search completed: May 12, 2006, 22:17:06  
Job time : 2384 secs

***This Page Blank (uspto)***



15	52.4	1.7	3330	11	US-10-9332-182A-84	Sequence 84, Appl	
16	52.4	1.7	3330	11	US-10-9332-182A-94	Sequence 84, Appl	
17	51.4	1.7	3304	13	US-10-979-095-19	Sequence 19, Appl	
18	48.6	1.6	3336	11	US-10-932-182A-75421	Sequence 75421, A	
19	48.6	1.6	3336	11	US-10-932-182A-75421	Sequence 75421, A	
20	47.8	1.6	1515	7	US-09-925-065A-5627	Sequence 5627, Ap	
21	47.8	1.6	1515	11	US-10-301-480-106864	Sequence 106864,	
22	47.8	1.6	1515	12	US-10-301-480-720273	Sequence 720273,	
23	45.8	1.5	1703	10	US-10-995-561-160	Sequence 160, App	
24	45.8	1.5	1949	10	US-10-995-561-161	Sequence 161, App	
25	45.8	1.5	1956	10	US-10-995-561-157	Sequence 157, App	
26	45.8	1.5	2781	10	US-10-995-561-285	Sequence 285, App	
27	45.8	1.5	25968	10	US-10-995-561-13248	Sequence 13248, A	
28	45.8	1.5	28933	10	US-10-995-561-13285	Sequence 13285, A	
29	45.2	1.5	2871	11	US-10-330-773-708	Sequence 708, App	
30	44.2	1.4	568	7	US-09-925-065A-119416	Sequence 119416,	
31	44.2	1.4	590	7	US-09-925-065A-508385	Sequence 508385,	
32	43.8	1.4	1095	11	US-10-932-182A-82043	Sequence 82043, A	
33	43.8	1.4	1095	11	US-10-932-182A-82043	Sequence 82043, A	
C	34	43.6	1.4	1331	17	US-11-136-527-2116	Sequence 2116, Ap
	35	43.6	1.4	1474	18	US-11-096-568A-16441	Sequence 16441, A
C	36	43.6	1.4	2261	18	US-11-051-720-72	Sequence 72, Appl
	37	43.6	1.4	2945	18	US-11-051-720-71	Sequence 71, Appl
38	43.6	1.4	6433	11	US-10-330-773-711	Sequence 711, App	
39	43.6	1.4	6927	13	US-10-960-414-231	Sequence 231, App	
40	43.6	1.4	7004	18	US-11-051-720-66	Sequence 66, Appl	
41	43.6	1.4	7128	18	US-11-051-720-65	Sequence 65, Appl	
42	43.6	1.4	7855	11	US-10-330-773-501	Sequence 501, App	
43	43.6	1.4	2077	17	US-11-136-527-548	Sequence 548, App	
44	43.4	1.4	2115	11	US-10-932-182A-3350	Sequence 3350, Ap	
45	43.4	1.4	2115	11	US-10-932-182A-3350	Sequence 3350, Ap	

```

43      43.4      1.4      2077      17      US-11-1308-3267-538      App
Sequence 348, App
44      43.4      1.4      2115      11      US-10-932-182A-3350      Ap
Sequence 3350, Ap
45      43.4      1.4      2115      11      US-10-932-182A-3350      Ap
Sequence 3350, Ap

ALIGNMENTS

RESULT 1
US-11-128-061-3433
; Sequence 3433, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128, 061

```

```

: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
:
: FILE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
:
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
:
: FILE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
:
: FILE REFERENCE: 01997.027701
:
: CURRENT APPLICATION NUMBER: US/11/128.061
:

```

```

; CURRENT FILING DATE: 2003-05-11
; PRIOR FILING NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3433
; LENGTH: 2703
; TYPE: DNA
; ORGANISM: Cricetus griseus
US-11-128-061-3433

Query Match          15.0%; Score 461.4; DB 17; Length 2703;
Best Local Similarity 61.5%; Pred. No. 1.1e-107;
Matches 779; Conservative 0; Mismatches 476; Indels 12; Gaps 2;

Qy 1226 GCGGCGCACAGGAGTGCACATCGCGGCCACCAACCCAGGAGCTGGCGGCGCTAGTACCGG 1285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 GGAGCGAGGAAGACGAAGTCCATCCCAACCGGACGACATCTTCTGCAAAACCTGCTCGGAG 600
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1286 GCGATCGCACGCAATTGAAGAGGTGGAGGCGACTGAAGAGGAGAAATCGCGAGAACGCGACGCGC 1345
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 601 GAAATCATGTATTGTGAAGAAATGAAAAAATGAAGAAACAAAGCGGGAAGAGAGGCG 660  
Qy 1346 CCGACAGCCGAGATGCAAGGAGAGAGAGTGGCGGTGATGACACAGGATCCGGGCAATCC 1405  
Db 661 CCAGAACTCGGAAATGAAGAAATGAAGCGAG- - - - -CTCAGGAATATGACAGCAGCTTTCC 714  
Qy 1406 AAATCGGAGAGCGGCGCAATGATACGCGAATATCAGACAGCTGGAATTTGTCCCGCT 1465  
Db 715 AAATCGGAAATTTGCCGATGATTAAGAAATTCGGGTACTATAGAAATGTCACCCACT 774  
Qy 1466 GTCGATGCCAGGCGTCGATGACCAATCAGATCAAGTGTGCGTGGCAAGCGTCCCAT 1525  
Db 775 TACTCTGACTGATCTACGGAAGAGCAGAGGATCTGCGTGTGTGAGGAGCGCCACT 834  
Qy 1526 TAGCGCAGAGGATCAATCGCAAGAGATCGATGTCATTTTCGGTGGCGCGCAAGGACAT 1585  
Db 835 GAATGAAGCAAGAACTGGCCCAAGAAAGAAATCGATGTGATTTCTGTTCGACGCAAGTGT 894  
Qy 1586 GCTCATCTGTCAGCAGCGCGCAGCAAGGTGCACTCAACAACTTCCTGGAGAACCAAA 1645  
Db 895 GCTCTTTGTGCATGAACCAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 954  
Qy 1646 GTTTCGCTTGCATAGCTTCAACGACAGCGTGGCAAGTGCATGATGATGATGATGATGAT 1705  
Db 955 ATTCTGCTTGCATTTGCTTGTATGAATAAAGCTTCAATGAAGTGTCTACAGGTTTCA 1014  
Qy 1706 AGCCAAGCGTTGTTGAACCAATTTTCGAGGCGGAATGGGAGCTGCTTCGCTACGG 1765  
Db 1015 AGCAAGGCACTGGTACAGAGATTTTGAAGGGGGGAAAGCAACTCTCTTTCGCTATGG 1074  
Qy 1766 CCAGACGGGATCGGGCAAAACCCACACATGGCGGTGAGTTTAAATGAAAGGTGCAGGA 1825  
Db 1075 GCAGACGGGAGTGGGAAGCTCACAAATGGGTGGAGACCTGTCTGTAAATCTCAGAA 1134  
Qy 1826 CTGAAGAAAGGCACTACGCAATGGCGGCCCAAGGATGCTTTGTGACCTTGAATATGCC 1885  
Db 1135 TACATCTAAAGGATCTATGCAATGGCTTCGCGGATGCTTCTCTGAGAGTGCAGCC 1194  
Qy 1886 GCGTTACCGGCCCATGATCTAGTCTCGGCCAGGTTTCTTGGAGATTTACAGTGGCA 1945  
Db 1195 TCGCTACCGGAACCTTAAACTTTGGAAGTTTATGTGACATTTCTTGAGATCTCAATGGGA 1254  
Qy 1946 GGTCTTCGATCTTCTGTCGCAACAGCAAACTCGCGCTCTCGGAGGATGTAACAGCA 2005  
Db 1255 GGTGTTGATCTGCTCAACAGAGGCCAAGTACTGATGATGATGATGATGATGATGATGAT 1314  
Qy 2006 AGTGAGGTGGGACTCAACGGAAGGTGGTGGATGGCGTTCGAGGAGGTACTGAAGCT 2065  
Db 1315 GGTGAGGTGGGACTTCAGGAGTACCTGGTTAACTGCTGATGATGATCATCAAGAT 1374  
Qy 2066 CATCCAGCACGCAATGTCGCCGAACATCCGCCAGAGCTGCGGCACTTCAATTCGTC 2125  
Db 1375 GCTCAACATGGGAGTGGCTGAGGACTTCTGGACAGACTTTTGGCAACTTCAATTTCTC 1434  
Qy 2126 GCGTTTCGACGCGCTTTTCCAGATTTGCTCGCGCGCGGCTCGAGGAAGATCCATGG 2185  
Db 1435 CCGCTCCCATGCTGCTTCCAGATTTCTTTCGAGCCAAAGG- - - - -GAGATTGATGG 1488  
Qy 2186 CAAGTTCTGTTCTATCGATCTGGCGGCAATGAGCGGGGCGTGGACACTTCTCTCGGCGGA 2245  
Db 1489 CAAATTTCTCTTGTGTGATCTGGCAGGGAATGAACAGAGGGGCTGACACTTCTAGTGCTGA 1548  
Qy 2246 TCGCAGACGGTATGGAGGGTCCGAGATTAACAAATTCGTCGTGCGCCCTCAAGGAGTG 2305  
Db 1549 CCGCCAGACTCGCATGGAGGGTGCAGAGATTAACAAAGTCTCTCTAGCACTGAAGGAGTG 1608  
Qy 2306 CATTCGTGCTGGGCAAAAGTGGGCCCACTTGCCTTCTGCTGCTCTCAACTCACCA 2365  
Db 1609 CATCAGGCGCTGGGACAGAACAGGCTCACACCCCATTTCTGAGAGCAAGCTGATCA 1668  
Qy 2366 GGTGCTGGCGACTGCTTTCATTTGGCGGAGAGAGCAAGCTGTCATGATGATGATGATGATGAT 2425

Db 1669 GGTCTAAGGANTTCATTTCATTTGGGAGAACTCGAGGACTTTCATGATTTGCCATGATCTC 1728  
Qy 2426 GCCGGGACTTTCCTTCGAGAGCAGCCTCAACACGCTGCGCTATGCGGATTCGTGTCAA 2485  
Db 1729 ACCAGGCAATAGCTCTCTGTAATATCTTTAAACACGCTGAGATATGCAGACAGATCAA 1788  
Qy 2486 GGAGCTG 2492  
Db 1789 GGAGCTG 1795  
RESULT 2  
US-11-128-049-3433  
; Sequence 3433, Application US/11128049  
; Publication No. US20060010513A1  
; GENERAL INFORMATION:  
; APPLICANT: Melville, Mark W.  
; APPLICANT: Charlebois, Timothy S.  
; APPLICANT: Mounts, William M.  
; APPLICANT: Hann, Louane E.  
; APPLICANT: Sinacore, Martin S.  
; APPLICANT: Leonard, Mark W.  
; APPLICANT: Brown, Eugene L.  
; APPLICANT: Miller, Christopher P.  
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR  
; TITLE OF INVENTION: MAKING AND USING SAME  
; FILE REFERENCE: 01997.027700  
; CURRENT APPLICATION NUMBER: US/11/128.049  
; CURRENT FILING DATE: 2005-05-11  
; PRIOR APPLICATION NUMBER: US 60/570,425  
; PRIOR FILING DATE: 2004-05-11  
; NUMBER OF SEQ ID NOS: 7285  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 3433  
; LENGTH: 2703  
; TYPE: DNA  
; ORGANISM: Cricetus griseus  
US-11-128-049-3433

Query Match 15.0%; Score 461.4; DB 17; Length 2703;  
Best Local Similarity 61.5%; Pred. No. 1.1e-107;  
Matches 779; Conservative 0; Mismatches 476; Indels 12; Gaps 2;  
Qy 1226 GGCGGCACAGGAGTCCGCTCGGCGCACCAACCCAGGAGCTGGCGGCTAGTACCCG 1285  
Db 541 GGAGGAGAGAGAGCAAGTCCATCCACCAGGACACATCTTCTGAAAACCTGCTCGAG 600  
Qy 1286 GCGATCCGACGCAATTTGAAGAGGTGGAGCGCACTGAAGAGAAATCGCGAGAGCGACGCG 1345  
Db 601 GAAATCATGTATTCTGAAGAAATGGAATAATGAAGAACAGCGGGAAGAGAGGCG 660  
Qy 1346 CCGACAGCGCAGATGAGAGGAGAGAGGTGGCGCTGATGAACAGGATCCGGGCAATCC 1405  
Db 661 CCAGAACTCGGAAATTAAGAAATAAGCGAG- - - - -CTCAGGAATATGACAGCAGCTTTCC 714  
Qy 1406 AAATCGGAGACGCGCAATGATACCGAATATCAGACAGCTGGAATTTGTGCGCT 1465  
Db 715 AAATCGGAAATTTGCCCGGATGATTAAGAAATTCGGGTACTATAGAAATGTCACCCACT 774  
Qy 1466 GCTCGATGGCGAGCGCTCGATGACCATCAGATGTCGCTGGCGCAAGCGTCCCAT 1525  
Db 775 TACTCTGACTGATCTTACGGAAGAGCAGATCTGCTGCTGTGAGGAAGCGCCACT 834  
Qy 1526 TAGCCGCAAGAGGTCAATCGCAAGAGATGATGATGATTTTCGGTGGCGCGCAAGGACAT 1585  
Db 835 GAATAAGCAAGAACTGGCCCAAGAAAGAAATCGATGTGATTTCTGTTCCAGCAAGTGT 894  
Qy 1586 GCTCATCTGTCAGAGCGCGCAGAGAGGTGCACTCACCAAGTTCCTCGAGAGACCAAA 1645  
Db 895 GCTCTTGTGATGAACCCCAATTTGAAGTGGACTTGAAGAGTATCTGGAGATATCAGGC 954  
Qy 1646 GTTTCGCTTCGATGAGCTTCAACGACAGCTGCGCAATGCCATGATGATGATGATGATGAT 1705

Db 955 ATTCTGCTCGATTTTGGCTTTGATGAAACAGCTTCAATGAAGTGTCTACAGGTTCCAC 1014  
Qy 1706 AGCCAAAGCGTGTGTGAAACCAATTTTCGAGGGCGGAATGCGAGCTGCTTCGCTACCG 1765  
Db 1015 AGCAAGGCCACTGTACAGACGATTTTGAAGGGGAAAGCAACCTGCTTGGCTATGG 1074  
Qy 1766 CCAGACGGGATCGGCAAAAGGCACACCATGCGGGGTGAGTTTAATGGAAGGTGCGGA 1825  
Db 1075 GCAGACGGGAGTGGGAAGACTCACAAATGGGGTGGAGACTGTGTGTTAAATCTCAGAA 1134  
Qy 1826 CTGCAAGAAACGGGACTCTACGCCATGCGGCGCAAGGATGCTTTGTGACCCCTGAATATGCC 1885  
Db 1135 TACATCTAAGGGATCTATGCAATGGCTTCCGGGATGTCTTCTCTGAAGAGTCAACC 1194  
Qy 1886 GCGTTACCGCGCCATGAATCTAGTCGTCTCGGCCAGTTTCTTTGAGATTTACAGTGGCAA 1945  
Db 1195 TCGCTACCGGAACTTAACTTGAAGTTTATGTGACATTTCTTTGAGATCTACAAATGGAA 1254  
Qy 1946 GGTCTTCGATCTTCTGTCGACAGCAAGAACTGCGGCTCTCGAGGATGGTAAACAGCA 2005  
Db 1255 GGTGTTGATCTGCTCAACAGAAAGGCCAAGCTACGTGTACTAGAAGACAGCAAGCAGCA 1314  
Qy 2006 AGTGCAAGTGGTGGGACTCACCGAGAAAGTGTGTCGATGGGTGCGAGGATCTGAAGCT 2065  
Db 1315 GGTGCAAGTGTGGGACTTCAGAGTACCTGTTTAACTGTGCTGATGATGTCAATGAT 1374  
Qy 2066 CATCCAGCAGCGCAATGCTGCCCAACATCCGCCAGAGCTCGGCCCAACTCCAAATTCGTC 2125  
Db 1375 GCTCAACATGGGAGTGCCTGCAGGACTTCTGGACAGACTTTTGGCCAACTCCAAATCTTC 1434  
Qy 2126 GGTTCGACGCGCTTTTCAGATTTGCTGCGGCGCAGGGCTCGACGAAAGATCCCATGG 2185  
Db 1435 CGCTCCCATGCTGCTTCCAGATTTCTTTCGAGCCAAAGG-----GAGATGTCATGG 1488  
Qy 2186 CAAGTTCTGCTTCATCGATCTGCGGCAATGAGCGGGGTGAGACATTTCTTCGCGCGCA 2245  
Db 1489 CAATTTCTCTGCTGTAATCTGACGGAATGAACGAGGGGCTGACACTTTCTAGTCTGA 1548  
Qy 2246 TCGGCGAGCGCTATGAGGGTCCGAGATTAACAAATCGCTGCTGCGCCCTCAAGGAGTG 2305  
Db 1549 CGCCAGACTCGCATGGAGGTGACAGATTAACAAAGTCTCTCTAGCAGTGAAGAGTG 1608  
Qy 2306 CATTCGTGCTGGGCAACAGTTCGGCCACTTGGCCCTTCGCTGTCTCCAACTCAACCA 2365  
Db 1609 CATCAGGGGCTGGGACAGAAAGGCTCACACCCCAATTCCTGAGAGCAAGCTGACTCA 1668  
Qy 2366 GGTGCTGCGGACTCGTTCAATTCGCGAGAGAGAGAGAGCTGTCATGATAGCCATGATCTC 2425  
Db 1669 GGTGCTAAGGGATTCATTCATTTGGGAGAACTCGAGACTTTCATGATTTGCCATGATCTC 1728  
Qy 2426 GCGGGACTTTCCTTCGAGCAGACAGCTCAACAGCTGCGCTATGCGGATCGTGTCAA 2485  
Db 1729 ACCAGGATTAAGCTCCTGTGAATATCTTTAAACAGCTGAGATATGCAGACAGATCAA 1788  
Qy 2486 GGAGCTG 2492  
Db 1789 GGAGCTG 1795

## RESULT 3

US-11-145-307A-22  
; Sequence 22, Application US/11145307A  
; Publication No. US20060094035A1  
; GENERAL INFORMATION:  
; APPLICANT: Acturus Bioscience, Inc.  
; APPLICANT: Erlander, Mark G.  
; APPLICANT: Ma, Xiao-Jun  
; TITLE OF INVENTION: Identification of Tumors  
; FILE REFERENCE: 022041-002020US  
; CURRENT APPLICATION NUMBER: US/11/145,307A  
; CURRENT FILING DATE: 2005-06-03  
; PRIOR APPLICATION NUMBER: US 60/577,084  
; PRIOR FILING DATE: 2004-06-04

; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 2832  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-145-307A-22

Query Match 14.7%; Score 453; DB 14; Length 2832;  
Best Local Similarity 62.5%; Pred. No. 1,6e-105;  
Matches 746; Conservative 0; Mismatches 435; Indels 12; Gaps 2;

Qy 1300 TGAAGAGGTGGAGCGACTGAAGGAGATCGCGAGAACGACGCGCCGACAGGCCGAGA 1359  
Db 647 TGAAGGAGTGGAAAAAATGAAGAAACGAGAGAGAGAAAGAGGCGCAGAACTCTGAAA 706  
Qy 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACAGGATCCGGGCAATCCAAACTGGGAGACGG 1419  
Db 707 TGAAGATGAAGAGAG-----CTCAGGAGTATGACAGTAGTTTTTCCAACTGGGAATTG 760  
Qy 1420 CGCAATGATACCGCAATATACAGACAGCTGCGAAATTTGTGCGCGCTGCTCGATGGCCAGG 1479  
Db 761 CCCGATGATTAAGAAATTTTCGGGCTACTTTGGAATGTCACTTACTATGACTGATC 820  
Qy 1480 CGGTGATGACCATCAGATCAGATGCTGTCGCGAAGCTCCATTTAGCCGCAAGAGG 1539  
Db 821 CTATCGAAGAGCAGACAGATATGTGTCTGTAGGAAACGCCCACTGAATAAGCAAGAT 880  
Qy 1540 TCATCGCAGGAGATCGATGTCATTTTCGGTGGCGGCGAGGACATGCTCATCTGTGACG 1599  
Db 881 TGCCCAAGAAAGAAATTTGATGTTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 940  
Qy 1600 AGCGCGCAGCAGAGGTGCGACCTCACCAAGTTCTCGAGAAACCAACAAAGTTTCGCTTCGACT 1659  
Db 941 AACCCTAAGTTGAAGTGGACTTAACAAAGTATCTGGAGAACCAAGCATTTCTGTTGACT 1000  
Qy 1660 AGCCTTCAACGACAGCTGCGACCAATGCCATGTGTATACAAATACAGCAAGCAAGCTGTG 1719  
Db 1001 TTGCATTTGATGAAACAGCTTCGAATGAAGTTGTCTACAGGTTTACAGCAGCAAGCCACTG 1060  
Qy 1720 TGAACACCAATTTTCGAGGGCGGAAATGGCGAGCTGCTTCGCTACGCGCAGACGGATCGG 1779  
Db 1061 TACAGCAATCTTTGAAGGTGGAAAGCAACTTGTTCATATGCGCAGCAGAGGAAGTG 1120  
Qy 1780 GCAAAACGACACCATGCGCGGTGAGTTTAAATGGAAGGTGACAGGACTGCAAGAACCGCA 1839  
Db 1121 GCAAGACACATATATGCGCGGAGACTCTCTGGGAAAGCCCAAGATGCAATCCAAAGGA 1180  
Qy 1840 TCTAGCCCATGCGCGCAAGGATGTCTTTGTGACCTGTAATATGCGCGGTACCGCGCCA 1899  
Db 1181 TCTATGCCATGGCTCCCGGACGCTTCTCTCTGAAGAAATCAACCTGTCTACCGGAAGT 1240  
Qy 1900 TGAATCTAGTCTCTCGCGCAGTTTCTTTGAGATTTACAGTTCGACAGGTTCTTCATCTTC 1959  
Db 1241 TGCGCTGGAAGTCTATGTGACATTTCTCGAGATCTCAATGGGAAGCTGTTTGAAGCTGC 1300  
Qy 1960 TGTCCGACAAAGCAGAAACTGCGCTCTCGAGGATGGTAAACAGCAAGTGCAGGTGTGG 2019  
Db 1301 TCAAAGAAAGGCCAAGCTGCGGCTGCTGGAGGACGCGAAGCAAGAGTGCAGTGTGG 1360  
Qy 2020 GACTACCCAGAGGTGTGTCGATGCGTCGAGAGGATCTGAAGCTCATCCAGCAGGCA 2079  
Db 1361 GCGCTGCAAGAACCTCTGGGCGAGACATTTGCCAACTTCCATTCCTCCCGCTCCACGCGT 1420  
Qy 2080 ATGCTGCCGCAACATCCGCGCAGAGCTCGGCCAACTCCAAATTCGTGCGTTCGACGCGG 2139  
Db 1421 GCGCTGCAAGAACCTCTGGGCGAGACATTTGCCAACTTCCATTCCTCCCGCTCCACGCGT 1480  
Qy 2140 TTTTCCAGATTTGCTGCGCGCGCAGGGCTCGACGAAAGATCCATGCAAGTTCTTCGTTC 2199  
Db 1481 GCTTCCAAATTAATTTCTTCAGAGCTAAAG-----GAGATGATGCAAGTCTCTTTGG 1534  
Qy 2200 TCGATCTGCGGCGCAATGAGCGGGCGTGGACACTTCTCTCGGCCGATCGGCGAGACGCGTA 2259

Db 1535 TAGATCGCAGGGAATGAGCGGCGGACACTTCCAGTGTGACCGGCAGACCGCA 1594  
Qy 2260 TGGAGGGTCCGAGATTAAACAAATCGCTGCGCCCTCAAGAGTGCATTTCGTGCTGG 2319  
Db 1595 TGGAGGGCCGAGAAATCAACAGAGTCTCTTAGCCCTGAAGAGGTGCATCAGGGCCCTGG 1654  
Qy 2320 GCAACAGTCGCGCCCACTTGCCTTCCCGTGTCTCCAAACTCACCCAGGTGCTGCGGACT 2379  
Db 1655 GACAGAACAAAGCTCACACCCGTTCCGTTGAGAGCAAGCTGACAGGTGCTGAGGACT 1714  
Qy 2380 CTTTCATTGGCGAGAGAGCAAGACGTGCATGATAGCCATGATCTGCGCGGACTTAGCT 2439  
Db 1715 CTTTCATTGGCGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCGGCAATAAGCT 1774  
Qy 2440 CTTGCGAGCACAGCTCAACAGCTGCGGTATGCGGATCGTTCAGGAGCTG 2492  
Db 1775 CCTGTGAATATACTTTAAACACCCCTGAGATATGCAGACAGGGTCAAGGAGCTG 1827

## RESULT 4

US-10-955-054A-137  
; Sequence 137, Application US/10955054A  
; Publication No. US20050266420A1  
; GENERAL INFORMATION:  
; APPLICANT: PUSZTAI, LAJOS  
; APPLICANT: SYMMANS, W. FRASER  
; APPLICANT: HESS, KENNETH R.  
; APPLICANT: AYERS, MARK  
; APPLICANT: STEC, JAMES  
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY  
; FILE REFERENCE: UTXC:880US  
; CURRENT APPLICATION NUMBER: US/10/955,054A  
; CURRENT FILING DATE: 2004-09-30  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 137  
; LENGTH: 2401  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-955-054A-137

Query Match 14.6%; Score 449.8; DB 10; Length 2401;  
Best Local Similarity 62.4%; Pred. No. 1e-104;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;  
Qy 1300 TGAAGAGCTGGAGCGACTGAAGGAGATCGGAGAGCGACGCGCCCGACAGCCGAGA 1359  
Db 770 TGAAGGAAGTGGAAAAATGAAGAAACAAGCGAGAGAGAAAGGCCCGCAAACTCTGAAA 829  
Qy 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACCGAGGATCCGGGCAATCCAACTGGGAGACGG 1419  
Db 830 TGAGAAATGAAGAGAG-CTCAGGAGTATGACAGTAGTCTTCCAACTGGGAAATTTG 883  
Qy 1420 CGCAATGATACGGGAATATCAGACAGCTGGAAATTTGTGCGCTGCTCGATGGCCAGG 1479  
Db 884 CCGGAATGATTAAGAAATTTCCGGCTACTTTGGATGTGATCCACTTACTATGACTGATC 943  
Qy 1480 CCGTGTGATGACCATCAGATCAGTGTGCGTGGCGGAGCGTCCCAATAGCCGCAAGGAGG 1539  
Db 944 CTATCGAAGAGCACAGAAATATGTGTGTGTAGGAAACGCGCCACTGAATAAGCAAGAAAT 1003  
Qy 1540 TCAATCGCAAGGAGATCGATGCTATTTCCGTGCGCGCAGGACATGCTCATCTGACG 1599  
Db 1004 TGGCAAGAAAGAAATGATGTGATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 1063  
Qy 1600 AGCCGCGCAGCAAGTTCGACCTTCACCAAGTTCTCTGGAGAACCAAGTTTCGCTTCGACT 1659  
Db 1064 AACCCAAAGTTGAAAGTGGACTTAAACAAAGTATCTGGAGAACCAAGCACTTCGCTTGA 1123  
Qy 1660 AGCCTTTCACGACACGTCGACATGCCATGCTGATATACAAATACACAGCCAGCGTTGG 1719  
Db 1124 TTGCAATTTGATGAAACAGCTTTTGAATGAAGTTGTCTACAGGTTTCACAGCAAGGCCACTGG 1183

Qy 1720 TGAAAACCATTTTCGAGGCGGGAATGGCGAGCTGCTTCCCTACGGCCAGAGCGGATCGG 1779  
Db 1184 TACAGCAATCTTTGAAGTGGAAAAGCAACTTGTGTTTTCATATATGGCCAGACAGGAATG 1243  
Qy 1780 GCAAAACGCACACCATCGGCGGTGAGTTTAATGAAAAGGTGCAGGACTGCAAGAACGGCA 1839  
Db 1244 GCAAGACACATACATATATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGGA 1303  
Qy 1840 TCTAGCCATCGGCGGCCAAGGATGTCTTTTGTGAACCTGAAATATGCGCGGTTCACGCGCCA 1899  
Db 1304 TCTATGCCATGGCTCTCCCGGACGCTCTTCTCTCTGAAAGAAATCAACCTCTGTACCGAAGT 1363  
Qy 1900 TGAATCTAGTGTCTCGCCAGTTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC 1959  
Db 1364 TGGGCTCTGGAAGTCTATGTGACATCTTTCGAGATCTCAATGGGAAGCTGTTTGAACCTGC 1423  
Qy 1960 TGTCCGACAAACGACAGAAACTCGGCGCTCTGGAGGATGGTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 1424 TCAACAGAAAGCCAGCTGCGCTGCTGGAGGACGGCAAGCAACAGGTGCAGTGGTGG 1483  
Qy 2020 GACTCACCGAAGGTGGTTCGATGGCGTFCGAGGAGGTACTGAAGTCAATCAGACCGCA 2079  
Db 1484 GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATCAAGATGCTCGACATGGGCA 1543  
Qy 2080 ATGCTGCCCGAACATCCGCCAGACGTCGGCCAACTCCAATTCGTCGCTTCGCAACGCG 2139  
Db 1544 CGGCTGCGAGAACTCTGGGCGAGACATTTGCAACTCCAATTCCTCCGCTCCCAACGCGT 1603  
Qy 2140 TTTTCCAGATTTGCTGCGCGCCGAGGCTCGAGAAAGATCCATGGGCAAGTTCTCGTTCA 2199  
Db 1604 GCTTCCAAATTTATTTTCGAGCTAAAGG-----GAGAAATGCATGGCAAGTTCTCTTGG 1657  
Qy 2200 TCGATCTGGCGGCAATGAGCGGGCGTGGACACTTCTCTCGGCCGATCGGCGAGCGGTA 2259  
Db 1658 TAGATCTGGCAGGGAATGAGCGAGCGCAGACACTTTCAGTGTCTGAACGCGCAGACCCGCA 1717  
Qy 2260 TGGAGGGTGGCGAGATTAAACAAATCGTGTGCGCCCTCAAGGAGTGCATTCGTGCGTGG 2319  
Db 1718 TGGAGGGCGCAGAAATCAACAAAGAGTCTTTAGCCCTGGAAGGATGCAATCAGGGCCCTGG 1777  
Qy 2320 GCAAAACAGTCGCGCCACTTTCGCTTCCAAACTCACCAGAGTGTGCGCGGACT 2379  
Db 1778 GACAGAACAGGCTCACACCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGAGT 1837  
Qy 2380 CGTTCATTGGCGAAGAGCAAGACGTCATGATGATAGCCATGATCTCGCCGGAATTAGCT 2439  
Db 1838 CTTTCATTGGGGAGAACTCTAGGACTTTCATGATTTGCCATGATCTCACAGGCAATAAGCT 1897  
Qy 2440 CTTGCGAGCACGCTCAACACGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492  
Db 1998 CCTGTGAATATACTTTTAAACACCCCTGAGATATGCAGACAGGGTCAAGAGCTG 1950

## RESULT 5

US-10-955-054A-101  
; Sequence 101, Application US/10955054A  
; Publication No. US20050266420A1  
; GENERAL INFORMATION:  
; APPLICANT: PUSZTAI, LAJOS  
; APPLICANT: SYMMANS, W. FRASER  
; APPLICANT: HESS, KENNETH R.  
; APPLICANT: AYERS, MARK  
; APPLICANT: STEC, JAMES  
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY  
; FILE REFERENCE: UTXC:880US  
; CURRENT APPLICATION NUMBER: US/10/955,054A  
; CURRENT FILING DATE: 2004-09-30  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 101  
; LENGTH: 2740  
; TYPE: DNA

ORGANISM: Homo sapiens  
US-10-955-054A-101

Query Match 14.6%; Score 449.8; DB 10; Length 2740;  
Best Local Similarity 62.4%; Pred. No. 1e-104;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;

```
QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGATCGGAGAGCGACGCCGCCGACAGCGCGAGA 1359
DB 638 TGAAGGAGTGGAAAAAATGAGAAACAAGCGAGAGAGAGAGAGCGCCAGAACTCTGAAA 697
QY 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACAGAGATCCGGCAATCCAAACTCGGAGACGG 1419
DB 698 TGAGATGAAGAG-----CTCAGGAGTATGACAGTAGTTTTTCCAACTGGGAATTTG 751
QY 1420 CGCAATGATACGCGAATATCAGAGCAGCTGCTGAATTTGTGCGCTGCTCGATGGCCAGG 1479
DB 752 CCCGAATGATTAAGAAATTTCCGGCTACTTTGGAATGTCTACCTACTTACTGATGATC 811
QY 1480 CCGTCGATGACCATCAGATCAGATGCTGCTGCGAAGCGTCCCATTTAGCCGCAAGGAGG 1539
DB 812 CTATCGAAGAGCAGAGATATGTCTGTGTTAGGAACGCCCACTGAATAAGCAAGAAT 871
QY 1540 TCAATCGCAGGAGATCGATGTCTATTTCCGCTGCGGCAAGGACATGCTCATCTGTCACG 1599
DB 872 TGCCCAAGAAAGAAATTTGATGATTTCCATTTCTAGCAAGTGTCTCTCTTTGGTACATG 931
QY 1600 AGCCGCGCAGCAAGGTGCGACTCAACAGTTCTCTGAGAAACCAAGTTTCGCTTCGACT 1659
DB 932 AACCCAAAGTTGAAAGTGGACTTAAAGAGTATCGGAGTATCGGAGTATCGGAGTATCG 991
QY 1660 AGCCCTTCAACGACAGCTGCGGAGTGGCGAGTATCGGAGTATCGGAGTATCGGAGTATCG 1719
DB 992 TTGCATTTGATGAACAGCTTCAAGTATGTTCTACAGTTTCTACAGTTTCTACAGTATCG 1051
QY 1720 TGAAGAACCATTTTCGAGGGCGGAATGCGGACGCTTTCGCTACGCGCAGCGGATCGG 1779
DB 1052 TACAGCAATCTTTGAGGTGGAAGAACTTTGTTTGCATATGCGCAGACAGGAGTG 1111
QY 1780 GCAAAACGCAACCATGCGGCTGAGTTTAAATGGAAGGTGCAAGGACTGCAAGAACGGCA 1839
DB 1112 GCAAGACACATACTATGGGCGGAGACTCTCTCGGAAAGGCCAGATGCTATCAAGGGA 1171
QY 1840 TCTACGCCATGCGCGCAAGGATGTCTTTGTGACCTGTAATATGCGCGGTACCGCGCCA 1899
DB 1172 TCTATGCCATGCGCTCCCGGAGCTTCTCTCTGGAAGTCAACCTGCTACCGGAACT 1231
QY 1900 TGAATCTAGTCTGCGGCGAGTTCTTTGAGATTTACAGTGGCAGAGTCTTCGATCTTC 1959
DB 1232 TGCGCCTGGAAGTCTATGTGACATCTTTCGAGATCTACAAATGGGAAGCTGTTTGACCTGC 1291
QY 1960 TGTCCGACAGCAGAAACTGCGGCTCTGAGGATGGTAAACAGCAAGTGCAAGTGGTGG 2019
DB 1292 TCAACAAGAGGCCAAGCTGGGCTGCTGGAGGACGCAAGCAAGTGCAAGTGGTGG 1351
QY 2020 GACTCACCGAAGAGTGGTGCATGCGGTGAGAGGATCTGAAGTCTATCCAGCAGCGCA 2079
DB 1352 GGTGCGAGGAGCATCTGCTTAACTCTGCTGATGATGCTATCAAGATGCTCGACATGGGCA 1411
QY 2080 ATGCTCGCCGACATCCGCGCAGAGCTGCGGCAACTCCAAATTCGTCGCTGCGAGCCG 2139
DB 1412 GCGCCTGCGAAGCTCTGGGAGACATTTGCCAACTCCAAATTCCTCCCGCTCCCAAGCT 1471
QY 2140 TTTTCCAGATTTGCTGCGCGCCGAGGCTCGACGAGAGTCCATGGCAAGTCTCTGCTTCA 2199
DB 1472 GCTTCCAAATTTCTTCGAGCTAAAG-----GAGATGATGATGCAAGTCTCTTTGG 1525
QY 2200 TCGATCTGCGGCAATGAGCGGGCGTGGACACTTCTCGGCGGATCGGCAAGCGGTA 2259
DB 1526 TAGATCTGCGGAGGATGAGCGAGGCGGAGACACTTCCAGTGTGAGCGGAGACCGCA 1585
QY 2260 TGGAGGCTCCGAGATTAACAACTGCTGCTGCGCTCAAGGATGCTATCGTGGCTTGG 2319
```

```
DB 1586 TGGAGGGCGCAGAAATCAACAGAGTCTTTAGCCCTGAAGGAGTGTGATCAGGGCCCTCG 1645
QY 2320 GCAAAACAGTCGCGCCACTTTGCCCTTCCCTGCTCTCCAACTCACCCAGGTGCTGCGCGACT 2379
DB 1646 GACAGAAACAGGCTCACACCCCGTTCCGTGAGAGCAAGCTGACACAGTCTCTGAGGACT 1705
QY 2380 CGTTCATTGCGGAGAGAGCAAGCTGCTGATGATAGCCATGATCTCGCCGGGACTTTAGCT 2439
DB 1706 CTTTCATTGCGGAGAACTCTAGGACTTGCATGATTCACAGATCTCACAGGCAAGACT 1765
QY 2440 CTTGCGGAGCAGCTCAACACAGCTGCTGCTATCGGATCGTGTCAAGGAGCTG 2492
DB 1766 CTTGCTGATATATCTTTAAACACCCCTGAGATATGACAGAGGCTCAAGAGCTG 1818
```

## RESULT 6

US-10-960-414-299  
; Sequence 299, Application US/10960414  
; Publication No. US20060074565A1  
; GENERAL INFORMATION:  
; APPLICANT: MILLER, LANCE D.  
; APPLICANT: GEORGE, JOSH  
; APPLICANT: VEGA, VINSENSIUS B.  
; TITLE OF INVENTION: METHODS, SYSTEMS, AND COMPOSITIONS FOR CLASSIFICATION,  
; TITLE OF INVENTION: PROGNOSIS, AND DIAGNOSIS OF CANCERS  
; FILE REFERENCE: 38271-76067  
; CURRENT APPLICATION NUMBER: US/10/960,414  
; CURRENT FILING DATE: 2004-10-06  
; NUMBER OF SEQ ID NOS: 500  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 299  
; LENGTH: 2740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-960-414-299

Query Match 14.6%; Score 449.8; DB 13; Length 2740;

Best Local Similarity 62.4%; Pred. No. 1e-104;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;

```
QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGATCGGAGAGCGACGCCGCCGACAGCGCGAGA 1359
DB 638 TGAAGGAGTGGAAAAAATGAGAAACAAGCGAGAGAGAGAGAGCGCCAGAACTCTGAAA 697
QY 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACAGAGATCCGGGCAATCCAAACTCGGAGACGG 1419
DB 698 TGAGATGAAGAG-----CTCAGGAGTATGACAGTAGTTTCCAACTGGGAATTTG 751
QY 1420 CGCAATGATACGCGAATATCAGAGCAGCTGGAATTTGTGCGCTGCTCGATGGCCAGG 1479
DB 752 CCCGAATGATTAAGAAATTTCCGGCTACTTTGGAATGTCTATCCACTTACTATGACTGATC 811
QY 1480 CCGTCGATGACCATCAGATCAGATGCTGCGTGGCAGAGCTGCCATTTAGCCGCAAGGAGG 1539
DB 812 CTATCGAAGAGCAGAGATATGTGCTGTGTAGGAAACGCCCACTGAATAAGCAAGAAT 871
QY 1540 TCAATCGCAAGGAGATCGATGTCTTTCCGTCGCGCAGGAGACATGCTCATCTGTCACG 1599
DB 872 TGCCCAAGAAAGAAATTTGATGATTTCCATTTCTAGCAAGTGTCTCTCTTTGGTACATG 931
QY 1600 AGCCGCGCAGCAAGGTGCGACTCAACAGTTCTCTGAGAAACCAAGTTTCGCTTCGACT 1659
DB 932 AACCCAAAGTTGAAAGTGGACTTAAAGAGTATCGGAGTATCGGAGTATCGGAGTATCG 991
QY 1660 AGCCCTTCAACGACAGCTGCGGAGTGGCGAGTATCGGATATGATGATGATGATGATGATG 1719
DB 992 TTGCATTTGATGAACAGCTTCAAGTATGTTCTACAGTTTCTACAGTTTCTACAGGCTTCA 1051
QY 1720 TGAAGAACCATTTTCGAGGGCGGAATGCGGAGTGGCGCTTTCGCTACGCGCAGAGCGGATCG 1779
DB 1052 TACAGCAATCTTTGAGGTGGAAGAACTTTGTTTGCATATGCGCAGACAGGAGTG 1111
QY 1780 GCAAAACGCAACCATGCGGCTGAGTTTAAATGGAAGGTGCAAGGACTGCAAGAACGGCA 1839
```

Db 1112 GCAAGACATATCTATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGA 1171  
Qy 1840 TCTAGCCATGGGGCCAAAGGATGCTTTGTGACCTGAAATATGGCGGTACCGGCCCA 1899  
Db 1172 TCTATGCCATGGCTCCCGGAGCGTCTTCTCTGGAAGATCAACCCCTACTACCGGAAGT 1231  
Qy 1900 TGAATCTAGTCTCTCGGCCAGTCTTTTGGAGATTTACAGTGGCAGGTCCTCGATCTTC 1959  
Db 1232 TGGGCTGGAAGTCTATGTGACATCTTCGAGATCTCAATGGGAGCTGTTTGACCTGC 1291  
Qy 1960 TGTCCGACAAAGCAGAAATCGCGCTCTGGAGGATGTTAAACAGCAAGTGCAGGTTGGTGG 2019  
Db 1292 TCAACAAGAAGCCAAAGCTGCGCTGTGGAGGACGGCAAGCAACAGGTGCAAGTGTGG 1351  
Qy 2020 GACTCACCAGAGAGTGTGATGGGCTGAGAGGTACTGAAGCTCATCCAGACGGCA 2079  
Db 1352 GGCTGGAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATCAAGATGCTCGACATGGGCA 1411  
Qy 2080 ATGCTGCCGGAACATCCGGCCAGACGCTCGGCCCAACTCCAAATTCGTCGGTTTCGACGCGG 2139  
Db 1412 GCGCTCGAAGACCTCTGGGAGACATTTGCCAACTCCAAATTCCTCCGCTCCCAAGGCT 1471  
Qy 2140 TTTTCCAGATGCTGCGGCCCGCAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTCA 2199  
Db 1472 GCTTCCAAATTTCTCTGAGCTAAAGG-----GAGAATGCATGGCAAGTTCTCTTTGG 1525  
Qy 2200 TCGATCTGGCGGCAATGAGCGGGGCTGGACACTTCTCGGCCCATCGGACAGCGTA 2259  
Db 1526 TAGATCTGGCAGGGAATGAGCGAGGCGCAGACACTTCCAGTCTGACCGGCGACACCGCA 1585  
Qy 2260 TGGAGGTCGCCAGATTAACAATCGCTGCTGGCCCTCAAGAGTGCAATCGTGGCTGG 2319  
Db 1586 TGGAGGCGCAGAAATCAACAAGATCTTTAGCCCTGAAGAGTGCAATCAGGCGCCCTGG 1645  
Qy 2320 GCAAAAGTTCGGCCCACTTGCCCTTCCTGTCTCAAAACTCACCAGGTGCTGCCGCACT 2379  
Db 1646 GACGAACAAGGCTCACACCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1705  
Qy 2380 GCTTCATTTGGCAGAGCAAGAGCTGATGATAGCATATTCGGCGGGACTTAGCT 2439  
Db 1706 CTTTCATTTGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCAGGCATAAGCT 1765  
Qy 2440 CTGGGAGCACAGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492  
Db 1766 CTTGGAATATCTTTAAACCCCTGAGATATGCACAGGCTCAAGGAGCTG 1818

RESULT 7  
US-11-128-061-7075  
; Sequence 7075, Application US/11128061  
; Publication No. US20060003958A1  
; GENERAL INFORMATION:  
; APPLICANT: Melville, Mark W.  
; APPLICANT: Charlebois, Timothy S.  
; APPLICANT: Mounts, William M.  
; APPLICANT: Hann, Louane E.  
; APPLICANT: Sinacore, Martin S.  
; APPLICANT: Leonard, Mark W.  
; APPLICANT: Brown, Eugene L.  
; APPLICANT: Miller, Christopher P.  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS  
; FILE REFERENCE: 01997.027701  
; CURRENT APPLICATION NUMBER: US/11/128,061  
; CURRENT FILING DATE: 2005-05-11  
; PRIOR FILING DATE: 2004-05-11  
; NUMBER OF SEQ ID NOS: 1400  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 7075  
; TYPE: DNA

; ORGANISM: Cricetulus griseus  
US-11-128-061-7075  
Query Match 7.9%; Score 243; DB 17; Length 1400;  
Best Local Similarity 69.6%; Pred. No. 9e-52;  
Matches 346; Conservative 0; Mismatches 145; Indels 6; Gaps 1;  
Qy 1996 GTAAACACCAAGTGCAGTGGTGGACATCACCGAGAGGTGTCATGGCGTCGAGGAGG 2055  
Db 2 GCAAGCAGCAGGTGCGAGTTGTGGACTTCAGGAGTACCTGCTTAACTGTGCTGATGATG 61  
Qy 2056 TACTGAAAGCTCATCCAGCACGCAATGCTGCCCAACATCCGGCCAGACGTCGGCCAACT 2115  
Db 62 TCATCAAGATGCTCAACATGGGCACTGCTCGAGACTTCTGACAGACTTTTGGCAACT 121  
Qy 2116 COAATTCGTCGCTTTCGACGCGGTTTTCAGATTTGTCGCGCCGAGGGCTCGACGA 2175  
Db 122 CCAATTTCTCCGCTCCCATGCTGCTTCCAGATTCTTCTTCGAGCCAAAGG-----GA 175  
Qy 2176 AGATCCATGGCAAGTTCCTGTTTCATCGATCTGGCGGGCAATGAGCGGGGCTGACACTT 2235  
Db 176 GATTGCATGGCAAAATTCCTTGGTGGATCTGGCAGGGAATGAACGAGGGCTGACACTT 235  
Qy 2236 CCTCGGCGATCGGCAGACGCGTATGAGGGGTGCCGAGATTAACAAAATCGCTGCTGGGCC 2295  
Db 236 CTAGTGTCTGACCGCAGACTCGCATGGAGGTGAGAGATTAAAGTCTCTCTAGACAC 295  
Qy 2296 TCAAGGAGTGCATTCGTGCTTGGGCAAAAGTCCGCGCCCACTTCCCTTCGCTGTCTCCA 2355  
Db 296 TGAAGGAGTGCATCAGGGCGCTGGGACAGAAACAGGCTCACACCCATTCCGTGAGAGCA 355  
Qy 2356 AACTCACCAGGTCTGCGGACTCGTTTCATTTGGGAGAGAGCAAGACGTGTCATGATAG 2415  
Db 356 AGCTGACTCAGGTGCTTAAGGGATTCATTTCATTGGGAGAACTCGAGGACTTGCATGTTG 415  
Qy 2416 CCATGATCTCCCGGAGCTTAGCTCCTGCGAGCACAGCTCAACAGCTGCGCTATGCGG 2475  
Db 416 CCATGATCTCACCAGGCATTAAGCTCCTGTGATATATCTTTAAACAGCTGAGATATGCG 475  
Qy 2476 ATCGTGTCAAGGAGCTG 2492  
Db 476 ACAGAGTCAAGGAGCTG 492

RESULT 8  
US-11-128-049-7075  
; Sequence 7075, Application US/11128049  
; Publication No. US20060010513A1  
; GENERAL INFORMATION:  
; APPLICANT: Melville, Mark W.  
; APPLICANT: Charlebois, Timothy S.  
; APPLICANT: Mounts, William M.  
; APPLICANT: Hann, Louane E.  
; APPLICANT: Sinacore, Martin S.  
; APPLICANT: Leonard, Mark W.  
; APPLICANT: Brown, Eugene L.  
; APPLICANT: Miller, Christopher P.  
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR  
; FILE REFERENCE: 01997.027700  
; CURRENT APPLICATION NUMBER: US/11/128,049  
; CURRENT FILING DATE: 2005-05-11  
; PRIOR FILING DATE: 2004-05-11  
; NUMBER OF SEQ ID NOS: 1400  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 7075  
; LENGTH: 1400  
; TYPE: DNA  
; ORGANISM: Cricetulus griseus  
US-11-128-049-7075  
Query Match 7.9%; Score 243; DB 17; Length 1400;







```
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOFUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; FILE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 201
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-201

Query Match      1.8%; Score 55.4; DB 18; Length 2217;
Best Local Similarity 49.0%; Pred. No. 0.0012;
Matches 270; Conservative 0; Mismatches 266; Indels 15; Gaps 4;

QY 1947 GTCTTCGATCTTCTGTCGGAAGAAGCAAGAACTCGCGTCTCTGGAGGATGTAACAGCAA 2006
    |||||
Db 1477 GGCATCTATGTTTCAGACCTCTCAACGACCTCTTCCTGCGCATCGAGGAGCCAGCAATGAC 1418

QY 2007 GTCCAGGTGTGGGACTCCCGAGAGAGGTGTCATGGCTGGCTCGAGGAGTACTGAAGCTC 2066
    |||||
Db 1417 ATGGAGTATGAGGTCTCCATGTCTTACCTGAGATCATGCAGCTGCTGATGAAGGGGAAC 1358

QY 2067 ATCCAGCAGCGCAATGTGCGCAACATCGGCGCAGACGTGGGCCCACTCCAATTCGTGCG 2126
    |||||
Db 1357 CGGCAGAGGACCCAGAGGCCACCGCGCCCAACAGACGTCTCCCGCTCCACGCGGTA 1298

QY 2127 CGTTCGACGCCGTTTTCAGATTTGCTGCGGCCCGAGGGCTCGACGAAGAT---CCAT 2183
    |||||
Db 1297 CTGCAGGTGACCGTGGCGCCAGCGCAGCGCGGTCAAGAACATCTTGCAGGAGGTGCGGCAG 1238

QY 2184 GGCAAGTTCTCGTTTCATGATCTGCGGGCAATGAGCGGGCGTGGACATCTCTCGGCC 2243
    |||||
Db 1237 GCGCGCTGTTTCATGATGACCTGGCTGAGCGCGCGCTCGCAGA---CACAGAT 1181

QY 2244 GATCGCAGACGCGTATGAGGGTGCAGAGATTAAACAAATCGCTGCTGCGCCCTCAAGGAG 2303
    |||||
Db 1180 CGTGGCAGCGTATGAAGGAGGGGCCACATCAACCGCTCACTGCTGGCATTGGGCAAC 1121

QY 2304 TGCAATTCGTGCGTTGGGCAACAGTCCGCCCACTTCCCTTC-----CGTGTCTCCAAA 2357
    |||||
Db 1120 TGCAATCAACGCCCTGAGCGCAAGGGTAGCAACAAAGTACATCAACTATCGCGACAGCAAG 1061

QY 2358 CTCACCCAGGTGTGGCGCACTCGTTTCATTTGGCGAGAGAGCAGAGCGTGCATGATAGCC 2417
    |||||
Db 1060 CTCACCCGCTCTTGAAGACTC---TCTGGGAGGAAACAGCGCAGTGCATGATCGCT 1004

QY 2418 ATGATCTCCCGGGACTTAGCTCTCGCAGACACACGCTCAACACGCTGCGCTATCGGAT 2477
    |||||
Db 1003 CACATCAGTCTCTGCGAGCAGTGCCTTCGAGGAGTCCCGGACACACCTTGACCTACGCCGC 944

QY 2478 CGTGTCAAGGA 2488
    |||||
Db 943 CGGGCCCAAGAA 933

RESULT 14
US-11-096-568A-4896
; Sequence 4896, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
```

```
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4896
; LENGTH: 1297
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1297)
; OTHER INFORMATION: Ceres Seq. ID no. 14304950
US-11-096-568A-4896

Query Match      1.8%; Score 54.6; DB 18; Length 1297;
Best Local Similarity 54.3%; Pred. No. 0.0017;
Matches 133; Conservative 0; Mismatches 109; Indels 3; Gaps 1;

QY 2258 TATGAGGGTCCGAGATTAAACAAATCGCTGCTGGCCCTCAAGGAGTGCATTCGTGCGTT 2317
    |||||
Db 46 TCTTGAGGGTGCATATATAACCGGTCTCTTCTTGCACTAAGCAGCTGTATTATTCTCT 105

QY 2318 GGGCAACAGTGGGCCACTTGGCCCTTCCGTGCTCCAACTCACCAGGTGCTGCGGA 2377
    |||||
Db 106 GGTAGAAGGCAAGAAACACATTCATCCATACCGAAACTCAAACTCACTCAACTTCTCAAGGA 165

QY 2378 CTCGTTTCATTTGGCGAGAGCAAGACGTCGATGATAGCATGATCTCGCCGGGACTTAG 2437
    |||||
Db 166 TTCACATA---GGAGGAACCTTGTAACTGTCTCATGATGTTCAACATAGCCCAAGTAACCT 222

QY 2438 CTCCTCGAGCAGCAGCTCAACACGCTGCGCTATCGGATCGTGTCAAGGAGCTGCTGGT 2497
    |||||
Db 223 CTCATTTGTTGAACCTCAGAACACAGTTCATTGGGCTGATAGACCAAGAGATTCCGGC 282

QY 2498 CAAGG 2502
    |||||
Db 283 AAAGG 287

RESULT 15
US-10-932-182A-84
; Sequence 84, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAWA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 84
; LENGTH: 3330
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-84

Query Match      1.7%; Score 52.4; DB 11; Length 3330;
Best Local Similarity 53.0%; Pred. No. 0.0074;
Matches 160; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

QY 2184 GGCAAGTTCTCGTTTCATCGATCTGGCGGCAATGAGCGGGCGTGGACACTTCTCTCGGCC 2243
    |||||
Db 934 GGTAAAGTTGAATTTGGTGGATTGGCTGGCAGTGA---GAACATTAAACAGATCAGGTGCG 990

QY 2244 GATCGGAGAGCGCTATCGAGGGTCCGAGATTAACAATCGCTGCTGCGCCCTCAAGGAG 2303
    |||||
Db 991 GAAACAAAAGAGCTCAAGAGCTGGAGTGAATTAACAAATCACTGTTGACTTTAGTCTGT 1050
```



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

QM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 17:52:51 ; Search time 530 Seconds  
(without alignments)  
10333.335 Million cell updates/sec

Title: US-10-840-060-142  
Perfect score: 3081  
Sequence: 1 aaactaaatgtgtgc.....gcgaatggcaagcgtagt 3081

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Issued Patents NA:\*
- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq:\*
  - 2: /cgn2\_6/prodata/1/ina/5 COMB.seq:\*
  - 3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*
  - 4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*
  - 5: /cgn2\_6/prodata/1/ina/H COMB.seq:\*
  - 6: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq:\*
  - 7: /cgn2\_6/prodata/1/ina/PP COMB.seq:\*
  - 8: /cgn2\_6/prodata/1/ina/RE COMB.seq:\*
  - 9: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	454.6	14.8	1292	3	US-09-594-669-1
2	454.6	14.8	1292	3	US-10-112-432-1
3	454.6	14.8	1420	3	US-09-594-669-7
4	454.6	14.8	1421	3	US-10-112-432-7
5	454.6	14.8	2172	3	US-09-594-669-13
6	454.6	14.8	2172	3	US-10-112-432-13
7	449.8	14.6	1175	3	US-09-594-669-3
8	449.8	14.6	1175	3	US-10-112-432-3
9	449.8	14.6	1304	3	US-09-594-669-9
10	449.8	14.6	1304	3	US-10-112-432-9
11	449.8	14.6	2740	3	US-09-594-669-15
12	449.8	14.6	2740	3	US-09-595-684B-32
13	449.8	14.6	2740	3	US-10-112-432-15
14	449.8	14.6	2740	3	US-09-849-602-5
15	425.6	13.8	1088	3	US-09-594-669-5
16	425.6	13.8	1088	3	US-10-112-432-5
17	425.6	13.8	1217	3	US-09-594-669-11
18	425.6	13.8	1217	3	US-10-112-432-11
19	412.2	13.4	417	3	US-09-270-767-2183
20	412.2	13.4	417	3	US-09-270-767-17465
21	404.4	13.1	702	3	US-09-270-767-31072
22	404.4	13.1	1774	3	US-09-270-767-14823
23	396.6	12.9	2505	3	US-09-595-684B-24
24	395.8	12.8	4617	3	US-09-620-312D-294

ALIGNMENTS

RESULT 1

US-09-594-669-1  
; Sequence 1, Application US/09594669  
; Patent No. 6331424  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6331424el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042  
; CURRENT APPLICATION NUMBER: US/09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: US 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1292  
; TYPE: DNA  
; ORGANISM: Human  
US-09-594-669-1

Query Match 14.8%; Score 454.6; DB 3; Length 1292;  
Best Local Similarity 62.6%; Pred. No. 6.2e-107;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

QY 1300 TGAAGAAGTGGAGCGATCGAAGAGATCGGAGAGCGAGCGCGCGAGCGCGAGA 1359  
Db 22 TGAAGGAAGTGGAAAAATGAAGACAAAGCGAGAGAGAGAGAGAGAGAGAGAG 81

QY 1360 TGAAGGAGGAGNAGTGGCGGTGATGAACAGGATCCGGGCAATCCAACTGGGACCG 1419  
Db 82 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTGTTCCTGGAATTTG 135

QY 1420 CGCAATGATACGCCAATATCAGACACGCTGGAATTTGTGCGGTGCTCGATGGCAGG 1479  
Db 136 CCGGAATGATTAAGAAATTTGGGCTACTTTGGAATGTCATCCATCTTACTGATGATC 195

QY 1480 CGGTGATGACCATCAGATCAGTGTGCGTGGCGAAGCGTCCCATTTAGCCGCAAGGAGG 1539  
Db 196 CTATCGAAGAGCAGAGATATGTGTCTGTGTGTAGGAAAGCGCCCACTGAATAAGCAAGAT 255

QY 1540 TCAATCGGAGGAGATCGATGTCATTTTCGGTCCCGCGGAGAGACATGCTCATGTCAGG 1599  
Db 256 TGGCAAGAAAGAAATTCATGTGATTTCCATTCCTAGCAAGGTCTCTCTTGGTACATG 315





Qy	2080	ATGTCGCCGGAACATCCCGGCAGACGTCGGCCAACTCCAAATTCGTGCGTTTCGACGCGC	2139
Db	796	GCGCCTGAGAAACCTCTCGGCAGACATTTGCGAACTCCAAATTCCTCCGCTCCCAACGCGT	855
Qy	2140	TTTTCCACATTGTCTGGCGCGCAGAGGCTCGACGAAGATCCATGGCAAGTTCTTCGTCA	2199
Db	856	GCTTCCAAATTTCTTCGAGCTTAAGG-----GAGAAATGCATGGCAAGTTCTCTTTGG	909
Qy	2200	TCGATCTGGCGGGCAATGAGCGGGCGTGGACACTTCTCGCCCGATCGGCAGACGCGTA	2259
Db	910	TAGATCTGGCAGGGCAATGAGCGAGGCGCGGACACTTCCAGTGTCTGACCGGCAGACCGCA	969
Qy	2260	TGGAGGGTCCGAGATTAACTGCTGTCTGGGCCCTCAGAGATGCAATTCGTGCGTTGG	2319
Db	970	TGGAGGGCGCAAAATCAACAAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGGCCCTGG	1029
Qy	2320	GCAACAGTCGGCGCCACTTGTGCCCTTCGATGTCCTCAAACTCACCCAGGTGTCGGCGACT	2379
Db	1030	GACAGAACAGGCTCACACCCGTTCCGTGAGCAAGCTGACACAGTGTCTAGGGACT	1089
Qy	2380	CGTTCTATGGCGAAGACGATGCAATGATAGCCATGATCTCCCGGGAGCTTAGCT	2439
Db	1090	CTTTCTATGGGAGAACTCTAGGACTTGCATGATGGCCAGATCTCACAGGCATTAAGCT	1149
Qy	2440	CCTCGGAGCACACGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTG	2492
Db	1150	CTGTGAAATATCTTTTAAACCCCTGAGATATGACAGAGGTCAAGGAGCTG	1202

RESULT 3  
US-09-594-669-7  
: Sequence 7, Application US/09594669  
: Patent No. 6331424  
: GENERAL INFORMATION:  
: APPLICANT: Beraud, Christophe  
: APPLICANT: Sakowicz, Roman  
: TITLE OF INVENTION: No. 6331424el motor proteins and methods for

```

? FILE REFERENCE: 1042
? CURRENT APPLICATION NUMBER: US/09/594,669
? CURRENT FILING DATE: 2000-06-15
? PRIOR APPLICATION NUMBER: US 09/295,612
? PRIOR FILING DATE: 1999-04-20
? PRIOR APPLICATION NUMBER: US 09/314,464
? PRIOR FILING DATE: 1999-05-18
? NUMBER OF SEQ ID NOS: 16
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 7
? LENGTH: 1420
? TYPE: DNA
? ORGANISM: Human
? US-09-594-669-7

```

[illegible]

RESULT 4  
US-10-112-432-7  
; Sequence 7, Application US/10112432  
; Patent No. 6638754  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6638754e1

Db	736	GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGGCA	795
Qy	2080	ATGCTGCCGCAACATCCGGCCAGACGCTCGGCCAACTCCAAATTCGTGCGCTTCGACGCGC	2139
Db	796	CGCGCTGCAGAACCTCTGGGACAGACATTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT	855
Qy	2140	TTTTCCAGATTGTGCTCGCGCCGACGGGCTCGACGAAAGATCCATGGCAAAGTTCTCGTTCA	2199
Db	856	GCITCCAAATTATCTTCGAGCTAAAG-----GAGATGCATGCGCAAGTTCTCTTTGG	909
Qy	2200	TCGATCTGGCGGGAATAGACGGGGCGTGGAACACTTCCTCGGCCGATCGGCAGACGCTA	2259
Db	910	TAGATCTGGCAGGAAATAGCAGGCGCGACACTTCCAGTGTCTGACCGGCAGACCCGCA	969
Qy	2260	TGGAGGTCGCGAGATTAACAATTCGCTGCTGCCCCCTCAAGAGTGCAATCGTGCCTTGG	2319
Db	970	TGAGGGCGCGAGAAATCAACAGAGTCTCTTAGCCCTGAAGGAGTGCATCAGGGCCCTGG	1029
Qy	2320	GCAAAAGTCGGGCCCACTTGGCCCTCCGTGCTCTCCAAACTCACCACGAGTGCTGCGCGACT	2379
Db	1030	GACAGAACAGGCTCACACCCCGTCCGTGAGAGCAAGCTGACACAGTGTCTGAGGACT	1089
Qy	2380	CGTTCAATTCGCGAAGAACAGACGCTGCATGATGCCATGATCGCCGGACTTAGCT	2439
Db	1090	CCITTCATTTGGGAGAACTCTAGGACTTGCATGATTGCCACGATCTCACCAGGCATAAGCT	1149
Qy	2440	CCTGCGAGCACGCTCAACACGCTGCGCTATCGCGATCGTCAAGAGCTG	2492
Db	1150	CTGTGAATATATCTTTAAACACCCCTGAGATATGCAGACAGGGTCAAGGAGCTG	1202
RESULT 5			
US-09-594-669-13			
; Sequence 13, Application US/09594669			
; Patent No. 6331424			
; GENERAL INFORMATION:			
; APPLICANT: Beraud, Christophe			
; APPLICANT: Sakowicz, Roman			
; TITLE OF INVENTION: No. 6331424el motor proteins and methods for			
; TITLE OF INVENTION: their use			
; FILE REFERENCE: 1042			
; CURRENT APPLICATION NUMBER: US/09/594,669			
; CURRENT FILING DATE: 2000-06-15			
; PRIOR APPLICATION NUMBER: US 09/295,612			
; PRIOR FILING DATE: 1999-04-20			
; PRIOR APPLICATION NUMBER: US 09/314,464			
; PRIOR FILING DATE: 1999-05-18			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 13			
; . LENGTH: 2172			
; TYPE: DNA			
; ORGANISM: Human			
US-09-594-669-13			
Query Match 14.8%; Score 454.6; DB 3; Length 2172;			
Best Local Similarity 62.6%; Pred. No.8e-107;			
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;			
Qy	1300	TGAAGAGGTGGAGCGACTGAAGGAGAAATCGCGAGAGCGACGCGCCGACAGCGCGAGA	1359
Db	578	TGAAGGAGTGGNAAAAATGAGACNACGCGAGAGAGNAGAGCCCAAGACTCTGAAA	637
Qy	1360	TGAAGGAGGAGAGGTGGCGCTGATGAACACAGGATCCGGGCAATCCAAACTGGGAGACGG	1419
Db	638	TGAGAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTTTTTCCAAACTGGGAATTG	691
Qy	1420	CGCAATGATACCGCAATATACGAGCACGCTGGAATTTGTCCGCTGCTCGATGCGCAGG	1479
Db	692	CCGGAATGATTAAGAAATTTCCGGGTACTTTTGGAAATGTCATCCACTTACTATGACTGATC	751
Qy	1480	CGCTCGATCACCATCAGATCAAGTGTGCGCAAGCGTCCCAATTTAGCCGCAAGGAGG	1539

Db 752 CTATCGAGACGACAGATATGTCTGTGTGTAGAAACGCCCTGATTAAGCAAGAAT 811  
Qy 1540 TCAATCGAAGGAGATCGATGTCAATTTGGTGGCGCGCAAGGACATGCTCATCTGTGCAAG 1599  
Db 812 TGGCCAAAGAAAGAAATGATGTGATTTCCATTTCTAGCAAGTGTCTCTTGTGTACATG 871  
Qy 1600 AGCCGCGCAGCAAGGTGACCTTCCACCAAGTTCCTGGGAACCAAGTTCGCTTCGACT 1659  
Db 872 AACCCTAAGTTGAAAGTGGACTTAACAAGATATCTGGGAACCAAGCAATTCGCTTTGACT 931  
Qy 1660 ACGCCTTCAACGACACGCTGCGACAAATGCCATGTGTATACAAATACACAGCCAAAGCGTTGG 1719  
Db 932 TTGCAATTTGATGAACAGCTTCGAATGAGTTGTCTACAGTTTCACAGCAAGGCCACTGG 991  
Qy 1720 TGAACAACATTTTCGAGGGCGGAATGCGAGCTGCTTCGCTACGGCCAGACGGGATCGG 1779  
Db 992 TACAGACAATCTTTGAAGGTGGAAGAACAACTGTTTTCATATGGCCAGACAGGAAGTG 1051  
Qy 1780 GCAAAACGACACCATCGGCGGTGAGTTTAATGGAAGGTGCAGGACTGCAGACGGCA 1839  
Db 1052 GCAAGACACATACTATGCGGGAGACCTCTCTGGGAAGGCCAGAAATGCAATGCCAAAGGGA 1111  
Qy 1840 TCTACGCGATGGCGGCCAAGATGCTTTGTGACCTGTGAATATATCGCGGTTCACCGGCCCA 1899  
Db 1112 TCTATGCCATGGCCTCCCGGGACGCTTCTCTGGAAGAACCCAGAAATGCAATGCCAAAGGGA 1171  
Qy 1900 TGAATCTAGTCTCGTGGCCAGTTCCTTTGAGATTTTACAGTGGCAAGGCTTCGATCTTC 1959  
Db 1172 TGGGCTTGGAAATCTATGTGACATCTTTCAGATCTTACAAATGGAAGCTGTTGACCTGC 1231  
Qy 1960 TGTCCGACAGCAGAACTCGGCTCTGAGAGTGTAAACAGCAAGTGCAGGTGGTG 2019  
Db 1232 TCAACAGAAGGCCAAGCTCGCGTGTGAGGATGGCAAGCAACAGGTGCAAGTGGTG 1291  
Qy 2020 GACTCACCGAGAGAGTGTGATGGCTGCGAGAGTACTGAACTCATCCAGCAGCGCA 2079  
Db 1292 GGCTGCAAGAGCATCTGTTAACTCTCTGATGATGTCATCAAGATGATCGACATGGGCA 1351  
Qy 2080 ATGCTGCCGCAACATCCGGCCAGACGCTGGCCCAACTCCAATTCGTCGCTTCGACGCGG 2139  
Db 1352 GCGCCTGCGAAGCTCTGGGCGAGACATTTGCCAACTCCAATTCCTCCGCTCCCAAGCT 1411  
Qy 2140 TTTTCCAGATTTGCTCGGCGCGAGGCTGCAAGATTCATGGAAGTTCCTGTTCA 2199  
Db 1412 GCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGAATGCAATGGAAGTTCCTTTGG 1465  
Qy 2200 TCGATCTGGCGGCAATAGCGGGGCTGACACTTCTCGGCGGATCGGCAAGCGTA 2259  
Db 1466 TAGATCTGGCAGGGAATGACGAGGCGCGGACACTTCCAGTGTGACCGGCGACCCGCA 1525  
Qy 2260 TGGAGGTGCGGAGATTAACAAATCGCTGTGGCCCTCAAGAGTGTGATTCGTGCGTTGG 2319  
Db 1526 TGGAGGCGCGAGAAATCAACAGAGTCTCTTAGCCCTGAGGAGTGCATCAGGGCCCTGG 1585  
Qy 2320 GCAACAGTCCGCCCATCTGCCCTTCTGCTGTCTCAAACTCAACCAAGTGTCTCGGACT 2379  
Db 1586 GACAGAACAGGCTCACAACCCGCTCGTGGAGCAAGCTGACACAGGTGCTGAGGACT 1645  
Qy 2380 CGTTTCAATGGCAGAGAGCAGACGCTGATGATAGCATGATCTCGCGGACTTAGCT 2439  
Db 1646 CTTTCATTTGGGAGAACTCTAGGACTTGCATGATGTCACAGATCTCACCAGGCGATAGCT 1705  
Qy 2440 CTTGCGAGCACGCTCAACACGCTGCTATGCGGATCTGTCAAGGAGCTG 2492  
Db 1706 CTTGTGATATATCTTTAAACACCCCTGAGATATGACAGAGGTCAAGGAGCTG 1758

RESULT 6  
US-10-112-432-13  
; Sequence 13, Application US/10112432  
; Patent No. 6638754  
; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6638754el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042A  
; CURRENT APPLICATION NUMBER: US/10/112,432  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: 09/724,215  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; PRIOR APPLICATION NUMBER: 09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 2172  
; TYPE: DNA  
; ORGANISM: Human  
US-10-112-432-13  
  
Query Match 14.8%; Score 454.6; DB 3; Length 2172;  
Best Local Similarity 62.8%; Pred. No. 8e-107;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
  
Qy 1300 TGAAGAGGTGGAGCGACTGACGAGAAATCGCGAGAGCGACGCGCCGACAGCCGAGA 1359  
Db 578 TGAAGAGGTGGAAAGAAATGMAACACAGCGAGAGAGAGAGGCCAGAACTCTGAAA 637  
Qy 1360 TGAAGAGAGAAAGTGGCGCTGTATGAACAGGATCCGGGCAATCCAAACTGGGAGACGG 1419  
Db 638 TGAGATGAAGAGAG-----CTCAGGAGTATGACAGTAGTTTTTCCAAACTGGGAATTG 691  
Qy 1420 CGCAATGTATACGCGAATATCAGAGCAGCGCTGGAATTTGTGCGCTGCTCGATGGCCAGG 1479  
Db 692 CCGCAATGTATTAAGAAATTTCCGGGCTACTTTGGAAATGTATCCACTTACTATGACTGATC 751  
Qy 1480 CCGTCCATGACCATCAGATCAGATGCTGTGCGCAAGCGTCCCATTTAGCCGCAAGGAGG 1539  
Db 752 CTATCGAAGAGCAGCAAGATATGTGTCTGTGTAGGAACGCCCACTGAATAAGCAAGAT 811  
Qy 1540 TCAATCGAAGGAGATCGATGTCAATTTCCGTGCGCGCAAGGACATGCTCATCGTGCACG 1599  
Db 812 TGGCCAAAGAAAGAAATTTGATGTGATTTCCATTTCTAGCAAGTGTCTCTCTTGTGTACATG 871  
Qy 1600 AGCCGCGCAGCAAGGTGCACTCACCAGTTCTCTGAGAACCAAGTTCGCTTCGACT 1659  
Db 872 AACCCTAAGTTGAAAGTGGACTTTAAACAAAGTATCTGGAGAACCAAGCATTTCTGCTTTGACT 931  
Qy 1660 ACGCCTTCAACGACACGCTGCGACAAATGCCATGTGTATACAAATACACAGCCAAAGCGTTGG 1719  
Db 932 TTGCAATTTGATGAACAGCTTCGAATGAAAGTTGTCTACAGTTTCACAGCAAGGCCACTGG 991  
Qy 1720 TGAACAACATTTTCGAGGGCGGAATGGCGACGTGCTTCGCTTACGGCCAGACGGGATCGG 1779  
Db 992 TACAGACAATCTTTGAAGGTGGAAAGCAACTTGTGTTTGCATATGCGCAGACAGGAAGTG 1051  
Qy 1780 GCAAAACGACACCATGCGCGGTGAGTTTAAATGGAAGGTGCAAGACTGCTCAAGAACGGCA 1839  
Db 1052 GCAAGACACATACTATGCGGGGAGACCTCTCTGGGAAGGCCAGAAATGCAATGCCAAAGGGA 1111  
Qy 1840 TCTACGCCATGGCGGCCAAGGATGCTTTGTGACCTGTGAATATATCGCGGTTCACCGGCCCA 1899  
Db 1112 TCTATGCCATGGCCTCCCGGGACGCTTCTCTCTGGAAGAAATCAACCTGCTACTCCGGAAGT 1171  
Qy 1900 TGAATCTAGTCTCGTGGCCAGTTCCTTTGAGATTTTACAGTGGCAAGGCTTCGATCTTC 1959  
Db 1172 TGGGCTTGGAAATCTATGTGACATCTTTCAGATCTTACAAATGGAAGCTGTTGACCTGC 1231  
Qy 1960 TGTCCGACAGCAGAAACTGCGCGCTCTCGAGGATGTTGTAACAGCAAGTGCAGGTGGTG 2019

Db 1232 TCAACAAGAGGCAAGCTGCGCTGCTGGAGATGGCAAGCAACAGAGTGCAAGTGGTG 1291  
Qy 2020 GACTCACCGAGAGGTGGTGGATGGCTGGAGAGGTACTGAAGCTCATCCAGCAGCGCA 2079  
Db 1292 GGTGAGAGGAGATCTGGTTAACTCTGCTGATGATGTCAATCAAGATGATCGACATGGGCA 1351  
Qy 2080 ATGCTGCCCGGAACATCCGCGCCAGAGAGTTCGGCCAACTCCAAATTCGTCCGCTTCGCACGCG 2139  
Db 1352 GCGCTGCAGAACCTCTGGGCGAGACATTTGGCCAACTCCAAATTCCTCCGCTCCAGCGGT 1411  
Qy 2140 TTTTCCAGATTTGCTGCGGCGCGAGGCTCGACGAAGATCATGGCAAGTTCTGTTCA 2199  
Db 1412 GCTTCCAAATTAATTTCTGAGCTAAAG-----GAGAATGATGCGCAAGTTCTCTTTGG 1465  
Qy 2200 TCGATCTGCGGCGCAATGAGCGGCGGTGGACACTTCTCGGCCATCGGCGAGACGGTA 2259  
Db 1466 TAGATCTGCGAGGGAATGAGCGGCGGACACTTCCAGTCTGACCGGCGAGACCCGCA 1525  
Qy 2260 TGGAGGCTGCCGAGATTAACAATCGCTGCTGGCCCTCAAGGAGTGCAATTCGTGGGTTGG 2319  
Db 1526 TGGAGGCGCGAGAAATCAACAGAGTCTTTAGCCCTGAAGAGTGCATCAGGCGCCTGG 1585  
Qy 2320 GCAACAGTGGCGCCACTTGGCCCTTCGCTGCTCCAACTCAACAGTGTGCGCGACT 2379  
Db 1586 GACAGAAACAAGGCTCACACCCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1645  
Qy 2380 CGTTCAATGGCGAGAGCAAGAGCTGCATGATAGCCATGATCTCGCGGGAAGTACT 2439  
Db 1646 CTTTCAATGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCGGCATAAGCT 1705  
Qy 2440 CTTGCGAGCACGCTCAACACGCTGCGCTATGCGGATCGTCTCAAGAGCTG 2492  
Db 1706 CTTGTGAATATCTTTAAACCCCTGAGATATGCGACAGGCTCAAGAGCTG 1758

## RESULT 7

US-09-594-669-3  
; Sequence 3, Application US/09594669  
; Patent No. 6331424  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6331424el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042  
; CURRENT APPLICATION NUMBER: US/09/594,669  
; CURRENT FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: US 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1175  
; TYPE: DNA  
; ORGANISM: Human  
US-09-594-669-3

Query Match 14.6%; Score 449.8; DB 3; Length 1175;  
Best Local Similarity 64.0%; Pred. No. 1e-105;  
Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

Qy 1402 ATCCAACTGGAGCGGCGGCAATGATACGGGAATATCAGAGCAGCGTGAATTTGTGC 1461  
Db 1 ATGCAAACTGGGAATTTGCCCGGAATGATTAAGAAATTTTCGGGCTACTTTGGAATGTGATC 60  
Qy 1462 CGCTCTCGATGCCGAGCGGTGATGACCATCAATCAGATCAGATGCGTGGCGCAAGGCTC 1521  
Db 61 CACTTACTATGACTGATCTTATCGAAGAGCAGAGATATGTTCTGTGTTAGGAACGCC 120  
Qy 1522 CCAATPAGCCGCAAGGAGTCAATCGAAGGAGATCGATGTCAATTTCCGTGCGCGCAAGG 1581

Db 121 CACTGAATTAAGCAAGAATTTGCCAAGAAAGAAATTTGATGTGATTTTCCATTTCTTAGCAAGT 180  
Qy 1582 ACATGCTCATCTGTGCAAGAGCGCGCAGCAAGGTGCGACTCACCAGTTTCTCTGAGAAACC 1641  
Db 181 GTCTCTCTTGTGATGAAACCAAGTTTGAAGTGGACTTAAACAAGTATCTGGAGAAC 240  
Qy 1642 ACAAGTTTTCGTCTCGACTTACGCTTCAACGACACGTGCGACAAATGCCATGCTATACAAAT 1701  
Db 241 AAGCAATCTGCTTTGACTTTGCAATTTGATGAACAGCTTCGAAATGAAGTTGTTCTACAGGT 300  
Qy 1702 ACACAGCCCAAGCCGTTGGTGAACCAATTTTCGAGGCGGGAATGGCGAGCTGCTTCGCGCT 1761  
Db 301 TCACAGCAAGGCCACTGGTACAGCAATCTTTTGAAGTGGAAAGCAACTTTGTTTTCAT 360  
Qy 1762 ACGCCAGACGGGATCGGCGAAACGACACCATCGGGGGTGAAGTTTAATGGAAGAGTGC 1821  
Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATATATGGGCGGAGACTCTCTCTGGAAAGCCC 420  
Qy 1822 AGGACTGCAAGAACCGGCATCTACGCCATGGCGGCAAGGATGCTTTTGTGACCCCTGAATA 1881  
Db 421 AAGATGCATCCAAAGGAGATCTATGCCATGGCTCCCGGAGCTCTTCTCTGAGAAATC 480  
Qy 1882 TGCCGCGTTACCGCGCCATGAATCTAGTCTGCTTCGCGCAGTTTCTTTGAGATTTACAGTG 1941  
Db 481 AACCTGCTACCGGAAGTTGGGCTTGGAAGTCTATGTGACATTTCTTCGAGATCTACAAATG 540  
Qy 1942 GCAAGGCTTTCGATCTTCTGTTCGCAAGCAGAAACTGCGGCTCTCTGAGGATGGTAAAC 2001  
Db 541 GGAAGCTGTTTGAACCTGCTCAACAAAGAGGCAAGCTGCGGCTGCTGGAGGATGGCAAGC 600  
Qy 2002 AGCAAGTGCAGGTGCTGGGACTCACCGCAAGAGTGGTTCGATGGGCTGCGAGGAGTACTGA 2061  
Db 601 AACAGGTGCAAGTGTGGGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATCA 660  
Qy 2062 AGCTCATCCAGCACGCGCAATGCTGCCCGAAACATCCGGCCAGACGCTCGGCCCAACTCCAATT 2121  
Db 661 AGATGATCGACATGGCGAGCGCTGCAGAACCTCTGCGGCGAGACATTTGCGCAACTCCAATT 720  
Qy 2122 CGTCGCTTGCACGCGCTTTTCCAGATTTGCTGCGGCGCGCAGGCTCGACGAGATCC 2181  
Db 721 CTTCCGCTCCACGCGTGTCTTCCAAATTAATTTCTCGAGCTTAAAG-----GAGAAATGC 774  
Qy 2182 ATGCAAGTTTCTGTTTCATCGATCTGCGGCGCAATGAGCGGCGTGGACACTTCTCTCGG 2241  
Db 775 ATGCAAGTTTCTTTTGTGATGCTGCGAGGGAATGAGCGGCGCGGACACTTCCAGTG 834  
Qy 2242 CCGATCGGCGAGACGCTATGAGGGTCCCGAGATTAACAAATCGCTCTGCGCCCTCAAGG 2301  
Db 835 CTGACCGCGAGACCCGATGAGGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGG 894  
Qy 2302 AGTCATTTGCTGCTTGGGCAACAGTCCGCGCACTTGGCCCTTCCGTGCTTCCAAACTCA 2361  
Db 895 AGTGCATCAGGCGCTTGGGACAGAAACAGGCTCACCCCGTTCCTGAGAGCAAGCTGA 954  
Qy 2362 CCAGGTGCTGCGGACTCTGTTTCATTTGCGGAGAGAGCAAGCTGATGATGCCATGA 2421  
Db 955 CACAGGTGCTGAGGACTCTTTCATTTGGGAGAACTCTTAGGACTTGCATGATGTCACGA 1014  
Qy 2422 TCTCGCGGGAATTTAGTCTCTCGAGCACACGCTCAACACGCTGCGCTATGCGGATCGTG 2481  
Db 1015 TCTCAACAGGATTAAGCTCTCTGTGAATATATCTTTAAACACCCCTGAGATATGACACAGGG 1074  
Qy 2482 TCAAGGAGCTG 2492  
Db 1075 TCAAGGAGCTG 1085

## RESULT 8

US-10-112-432-3  
; Sequence 3, Application US/10112432  
; Patent No. 6638754  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe

APPLICANT: Sakowicz, Roman  
TITLE OF INVENTION: No. 631754el motor proteins and methods for  
TITLE OF INVENTION: their use

FILE REFERENCE: 1042A

CURRENT APPLICATION NUMBER: US/10/112.432

CURRENT FILING DATE: 2002-03-29

PRIOR APPLICATION NUMBER: 09/724,215

PRIOR FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: 09/295,612

PRIOR FILING DATE: 1999-04-20

PRIOR APPLICATION NUMBER: 09/314,464

PRIOR FILING DATE: 1999-05-18

PRIOR APPLICATION NUMBER: 09/594,669

PRIOR FILING DATE: 2000-06-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 1175

TYPE: DNA

ORGANISM: Human

US-10-112-432-3

Query Match 14.6%; Score 449.8; DB 3; Length 1175;

Best Local Similarity 64.0%; Pred. No. 1e-105;

Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

QY 1402 ATCAAACTGGGAGAGCGCGCAATGATACGCGAATATACAGACAGCTGGAAATTTGTGC 1461

Db 1 ATGCAAACTGGGAATTTGGCCGAATGATTAAGAAATTTTCGGGTACTTTGGAAATGTCAATC 60

QY 1462 CGCTGCTCGATGCCAGCGCTCGATGACCATCAGATCAGATGCTGCGTGGCGCAAGCGTC 1521

Db 61 CACTTACTATGATCTGATCTTATCGAAGAGCACAGAATATGTGTGTGTGTAGGAACGCC 120

QY 1522 CCAATTAGCCCAAGGAGGTCAATCGCAAGGAGATCGATGTCATTTTCGGTCCCGCGCAAGG 1581

Db 121 CACTGAATAGCAAGAAATGGCCCAAGAAAGAAATTTGATGATTTCCATTTCTAGCAAGT 180

QY 1582 ACATGCTCATCTGTCAGAGCGCGGAGCGGAGTGCACCTTACCAAGTTCTCTGGAGAAC 1641

Db 181 GTCTCTCTTGTGTACATGAACCAAGTTGAAAGTGGACTTAAACAAAGTATCTGGAGAAC 240

QY 1642 ACAAGTTTCGTTTCGACTAGCGCTTCAACGACAGTGGCAGATGCGATGATGATACAAAT 1701

Db 241 AAGCATTTCTTTGATCTTTCATTTGATGAACAGCTTTCGAATGAAGTTGTCTACAGGT 300

QY 1702 ACACAGCCCAAGCGTTTGTGAAACCAATTTTCAGGCGCGAATGGCGAGCTGCTTCGCGCT 1761

Db 301 TCACAGCAAGGCCACTTGTACAGACATCTTTGAGGTGGAAAGCAACTTTGTTTGCAT 360

QY 1762 ACGCCGAGAGCGGATCGGGCAAAACGCACACCATGGCGGTGAGTTTAAATGGAAGAGTGC 1821

Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATACTATGGCGGAGACCTCTCTGGGAAGCCC 420

QY 1822 AGGACTGCAAGACGCGATCTAGCCATGGCGCCCAAGATGCTTTTGTGACCTGAATA 1881

Db 421 AGAATGCATCAAGGGATCTATGSCATGGCTCCCGGAGCGTCTTCTCTGGAAGATC 480

QY 1882 TGCCTGGTTACCGCGCCCAATGATCTAGTCTCTCGGCCAGTTTCTTTTGAGATTTACAGTG 1941

Db 481 AACCCTGCTACCGGAAGTTGGGCTGGAAGTCTATGTGACATCTTCGAGATCTACATG 540

QY 1942 GCAAGTCTTCGATCTTCTGTCCGACAGCAAACTGGCGCTCTGGAGGATGGTAAC 2001

Db 541 GGAAGCTGTTGACCTGCTCAACAGAGGCCAAGCTGGCGGTCTGGAGGATGGCAAGC 600

QY 2002 AGCAATGTCAGTGTGGGACTCACCGAAGAGTGTGATGGCGTCTGGAGGAGTACTGA 2061

Db 601 AACAGTGCAGTGTGGGCTGTCAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATCA 660

QY 2062 AGCTCATCCAGCAGCAATGTCTGCCGCAACATCCGCGCAGAGCTCGGCCCACTCAAT 2121

Db 661 AGATGATCGACATGGGCAAGCGCTGCAAGAACTCTGGGAGACATTTTGCCTCAAT 720

QY 2122 CGTCCGCTTCGACAGCGGTTTTCAGATTGTCTCGCGCGCAGGGCTCGACGAAGATCC 2181

Db 721 CTTCCCGCTCCACACGGTGTCTTCCAAATATTCTTCGAGCTAAAG-----GAGAATGC 774

QY 2182 ATGGCAAGTCTCGTTTCATCGATCTGGCGGCAATGAGCGGGCGTGGACACTTCTCTCG 2241

Db 775 ATGGCAAGTCTCTTTTGGTAGATCTGGCAGGGAATGAGCGAGCGCGGACACTTCCAGTG 834

QY 2242 CCGATCGGACAGCGCGTATATGGAGGGTCCGAGATTAAACAAATCGTCTGGCCCTCAAGG 2301

Db 835 CTGACCGGACAGCCCGCATGGAGGGCGCAGAAATCAACAGAGTCTCTTAGCCCTGAGG 894

QY 2302 AGTGCAATCTGCTGTTGGGCAACAGTCCGCCACCTTTCGGTGTCTCCAAACTCA 2361

Db 895 AGTGCAATCAGGGCCCTTGGGACAGAACAGGCTCACACCCCGTTCCTCGTAGAGCAAGCTGA 954

QY 2362 CCAGGTGCTGCGGACTCGTTTCATTTGGCGAGAGAGCAAGCTGTCATGATAGCCATGA 2421

Db 955 CACAGTGTCTGAGGGACTCTTTCATTTGGGAGAACTCTTAGGACTTGCATGATTTGCCACGA 1014

QY 2422 TCTCGCCGGACTTAGTCTCTCGAGCACACGCTCAACAGCTGCGCTATGCGGATCGTG 2481

Db 1015 TCTCACCAGGCATAGCTCTCTGTGAATATATCTTTAAACACCTTGAGATATGACAGACGG 1074

QY 2482 TCAAGGAGCTG 2492

Db 1075 TCAAGGAGCTG 1085

#### RESULT 9

US-09-594-669-9

Sequence 9, Application US/09594669

Patent No. 6331424

GENERAL INFORMATION:

APPLICANT: Beraud, Christophe

APPLICANT: Sakowicz, Roman

TITLE OF INVENTION: No. 6331424el motor proteins and methods for

TITLE OF INVENTION: their use

FILE REFERENCE: 1042

CURRENT APPLICATION NUMBER: US/09/594,669

CURRENT FILING DATE: 2000-06-15

PRIOR APPLICATION NUMBER: US 09/295,612

PRIOR FILING DATE: 1999-04-20

PRIOR APPLICATION NUMBER: US 09/314,464

PRIOR FILING DATE: 1999-05-18

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9

LENGTH: 1304

TYPE: DNA

ORGANISM: Human

US-09-594-669-9

Query Match 14.6%; Score 449.8; DB 3; Length 1304;

Best Local Similarity 64.0%; Pred. No. 1.1e-105;

Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

QY 1402 ATCAAACTGGGAGAGCGCGCAATGATACGCGAATATACAGACAGCTGGAAATTTGTGC 1461

Db 1 ATGCAAACTGGGAATTTGGCCGAATGATTAAGAAATTTTCGGGTACTTTGGAAATGTCAATC 60

QY 1462 CGCTGCTCGATGGCGCGCTCGATGACCATCAGATCACAGTGTGCGTGGCGCAAGCGTC 1521

Db 61 CACTTACTATGATCTGATCTTATCGAAGAGCACAGAATATGTGTGTGTAGGAACGCC 120

QY 1522 CAATTAGCGCGAAGGAGTCAATCCGAAGGAGATCGATGCTATTTTGGTGGCGCGCAAGG 1581

Db 121 CACTGAATAAGCAAGAAATTTGGCCAAAGAAAGAAATTTGATGTGATTTCCATTTCTTAGCAAGT 180

QY 1582 ACATGCTCATCTGTCAGCGCGCGCAGCAAGTGCACCTCACCAGTTCCTTGGAGAAC 1641

Db 181 GTCTCTCTTGTGTACATGAACCCAAAGTTGAAAGTGGACTTAAACAAAGTATCTGGAGAAC 240

```
QY 1642 ACAAGTTTCGCTTCAAGTACGCTTCAACGACACGTCGACAAATGCCATGCTATACAAAT 1701
Db 241 AAGCATTTCTGCTTTGACTTTGACTTTGATGAAACACAGCTTCGAATGAAGTTGTACAGT 300
QY 1702 ACACAGCAAGCGTTTGGTGAACCAATTTTCAGAGCGGAATGCGAGCTGCTTCGCT 1761
Db 301 TCACAGCAAGCGCTTGGTACAGCAATCTTTGAAGTGGAAAGCAACTTGTGTTGCAAT 360
QY 1762 ACGGCCAGAGGATCGGCCAAACCCACACCATGGCGGTGAGTTTAATGAAAGGTGC 1821
Db 361 ATGGCCAGACAGGAAGTGCAAGACACATACCTATATGGCGGAGACCTCTCTGGGAAAGCCC 420
QY 1822 AGGACTGCAAGAACGCACTTACGCCATGGCGGCCAAGGATGCTTTGTGACCTCGAATA 1881
Db 421 AGAATGCATCCAAAGGATCTATGCATGGCCCTCCGGGAGCTCTCTCTCTGAAGATC 480
QY 1882 TGCCGGTTACCGCCCATGAATCTAGTGTCTCGGCCGATTTCTTTGAGATTTACAGTG 1941
Db 481 AACCCCTGCTACCGGAAGTTGGGCCCTGGAAGTCTATGTGACATTTCTTCGAGATCTACAATG 540
QY 1942 GCAAGCTCTTCGATCTTCTGTCGACAGCAAGCAAACTCGGCTCTCGAGGATGTTAAC 2001
Db 541 GGAAGCTGTTGACCTGCTCAACAAAGAGGCCCAAGCTGCGCGTCTGGAGGATGGCAAGC 600
QY 2002 AGCAAGTGCAGGTGCTGGGACTCACCGAGAAAGGTGCTGATGGCGTCTGAGGAGTACTGA 2061
Db 601 AACAGTGCAGTGTGGGCTGACGAGCATCTGTTAACTCTGCTGATGATGTCATCA 660
QY 2062 AGCTCATCAGCA CGGCAATGCTGCCCGAACAATCCGGCAGACGTCGCGCCAACTCCAATT 2121
Db 661 AGATGATCGACATGGGCAGCGCTGCAGAACTCTGCGGACAGACATTTGCCAACTCCAATT 720
QY 2122 CGTCGCTTCGACGCGCTTTTCCAGATTTGCTGCGCGCGGAGGCTCGACGAGATCC 2181
Db 721 CCTCCGCTCCACGCGCTCTCCAAATTAATTTCTCGAGCTAAAGG-----GAGATGC 774
QY 2182 ATGGCAAGTCTCTGTTTCATCGATCTGGCGGCAATGAGCGGCGTGACACTTCTCTCGG 2241
Db 775 ATGGCAAGTCTCTTGTGTAGATCTGGCAGGGAATGAGCGGCGGACACTTCCAGTG 834
QY 2242 CCGATCGCAGACGCGTATGAGGGTGCAGAGATTAACAAATCGCTGCTGCGCTTCAAGG 2301
Db 835 CTGACCGGACAGACCGCATGAGGCGCAGAAATCAACAAGATCTCTTAGCCCTGGAAG 894
QY 2302 AGTGCAATTCGTGCTGGTGAACAGTGGCGCCACTTCCGCTTCCGTTCTCCAAACTCA 2361
Db 895 AGTGATCAGGCGCTGGGACAGAAACAAAGGCTCACACCCGCTTCGCTGAGAGCAAGCTGA 954
QY 2362 CCCAGTCTCGCGACTCGTTTCATTGGCGAGAGAGCAAGCTGCAATGATAGCATGA 2421
Db 955 CACAGTCTGAGGACTTCCTTCAATGGGAGAACTCTAGGACTTGCATGATTTGCCAGA 1014
QY 2422 TCTCGCGGACTTAGCTCTGCGAGCACAGCTCAACACGCTGCGCTATCGGATCGTG 2481
Db 1015 TCTACCAAGGCATAAGCTCTCTGTAATATCTTTAAACACCCCTGAGATATCAGACAGGG 1074
QY 2482 TCAAGGAGCTG 2492
Db 1075 TCAAGGAGCTG 1085
```

## RESULT 10

```
US-10-112-432-9
; Sequence 9, Application US/10112432
; Patent No. 6638754
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6638754e1 motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/112.432
```

```
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Human
US-10-112-432-9
```

```
Query Match 14.6%; Score 449.8; DB 3; Length 1304;
Best Local Similarity 64.0%; Pred. No. 1.1e-105;
Matches 598; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

QY 1402 ATCCAACTGGGAGCGCGCAAAATGATACCGAATATCAGAGCACGCTGGAATTTGTGC 1461
Db 1 ATGCAAACTGGGAATTTGCCGAATGATTAAAGAAATTTCCGGCTACTTTTGGAAATGTCATC 60
QY 1462 CGCTGCTCGATGGCCAGCGCTCGATGACCATCAGATCACAGTGTGCGTGGCAGCGTC 1521
Db 61 CACTTACTATGACTGATCTCTATCGAAGAGCAGCAATATGTGTCTGTGTTAGGAAACGCC 120
QY 1522 CCATTAGCCGCAAGAGGTCAATCGAAGGAGATCGATGTCATTTCCGTCGCGCAAGG 1581
Db 121 CACTGAATACCAAGAAATTTGCCAAGAAATTTGATGATTTTCCATTTCTAGCAAGT 180
QY 1582 ACATGCTCATCGTGCA CGAGCGCGCAGCAAGGTGCGACCTCACCAGTTCCTGGAGAAC 1641
Db 181 GTCTCCTCTTGGTACATGAACTTGAAGTGGACTTAACAAGATATCTGGAGAAC 240
QY 1642 ACAAGTTTCGCTTTCGACTACGCTTCAACGACAGTGGGACAAATGCCATGTTATCAAT 1701
Db 241 AAGCATTTCTGCTTTGACTTTTGCATTTTGAATGAAACAGCTTTCGAATGAAATGCTTACAGGT 300
QY 1702 ACACAGCCAAAGCGCTTGTGTAACCAATTTTCGAGGCGGAAATGCGCAGCTTCTCGCT 1761
Db 301 TCACAGCAAGGCCACTGTTACAGACAACTTTTGAAGTGGAAAGCAACTTTGTTTGGAT 360
QY 1762 ACGGCCAGACGGGATCGGGCAAAACCGCACCATGCGGCGGTGAGTTTAATGGAAGGTGC 1821
Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATATATGGCGGAGACCTCTCTGGGAAAGCCC 420
QY 1822 AGGACTGCAGAAACGGGCTCTACGCCATGGCGGCAAGGATGCTTTGTGACCTGATTA 1881
Db 421 AGAATGCATCCAAAGGATCTATGCCATGGGCTCCCGGAGCTCTTCTCTGGAAGATC 480
QY 1882 TGCCGCTTACCGCGCCATGAATCTAGTCTGCTCGGCGAGTTTCTTTGAGATTTACAGTG 1941
Db 481 AACCTCTACCGGAAGTTGGGCTTGGAACTCTATGTGACATTTCTTCGAGATCTACAATG 540
QY 1942 GCAAGGTCTTCGATCTTTCTGTCCGACAGCAAGAACTGCGCGTCTCTGAGGATGGTAAAC 2001
Db 541 GGAAGCTGTTTGACCTCTCAACAAAGAGGCCAAAGCTGCGCGTCTGCGAGGATGGCAAGC 600
QY 2002 AGCAAGTGCAGGTGCTGGGACTCACCGAGAGGTTGTCGATGGCGTGGAGGAGTACTGA 2061
Db 601 AACAGGTCAGATGTTGGGCTGCGAGGATCTTGGTTAACTCTGCTGATGATGTCATCA 660
QY 2062 AGCTCATCCAGCACGCGCAATGCTGCCGCAACATCCGGCCAGACCTCGGCGCAACTCCAATT 2121
Db 661 AGATGATCGACATGGGCGAGCGCTGCAAGACCTCTGGGCGAGACATTTGCCAACTCCAATT 720
QY 2122 CGTCGCTTCGACGCGCTTTTTCAGATTTGCTGCGCGCGGAGGCTCGACGAGATCC 2181
Db 721 CCTCCGCTCCACGCGTCTTCCAAATTAATTTCTCGAGCTAAAGG-----GAGATGC 774
```



QY 2182 ATGCAAGTCTCTGTTTCATCGATCTGCGGGCAATGAGCGGGGGTGGACATCTTCTCGG 2241  
Db |||||  
775 ATGCAAGTCTCTTCTTGGTAGATCTGGCAGGAATGAGCGGGCGGACATCTCCAGTG 834  
QY 2242 CCGATCGGAGACCGGTATGAGGGTCCGAGATTAAACAATCGCTCGCCCTCAAGG 2301  
Db |||||  
835 CTGACCGGACAGCCCGATGGAGGGCGAGAAATCAACAAGTCTCTTAGCCCTGAGG 894  
QY 2302 AGTGCAATCTGCGTTGGGCAACAGTCCGCCCACTTGGCCCTTCCGTTCTCCAAACTCA 2361  
Db |||||  
895 AGTGCAATCAAGGGCCCTGGGACAGAAACAAGCTCACACCCGTTCCGTCGAGCAAGCTGA 954  
QY 2362 CCGAGGTGTCGCGGACTCTCTTCAATGGCGAGAGCAAGCTGATGATAGCCATGA 2421  
Db |||||  
955 CACAGGTGCTGAGGGACTCTCTTCAATGGGAGAACTCTAGGACTTGCATGATGCCACGA 1014  
QY 2422 TCTCGCGGACTTAGTCTCTCGGAGCACACGCTCAACAGCTGCGCTATGCGGATCGTG 2481  
Db |||||  
1015 TCTCACCGGCATAAGCTCTCTGATATATCTTTAAACACCCCTGAGATGCAAGAGG 1074  
QY 2482 TCAAGGAGCTG 2492  
Db |||||  
1075 TCAAGGAGCTG 1085

## RESULT 11

US-09-594-669-15

; Sequence 15, Application US/09594669

; Patent No. 6331424

; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe

; APPLICANT: Sakowicz, Roman

; TITLE OF INVENTION: No. 6331424el motor proteins and methods for

; FILE OF INVENTION: their use

; FILE REFERENCE: 1042

; CURRENT APPLICATION NUMBER: US/09/594,669

; CURRENT FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: US 09/295,612

; PRIOR FILING DATE: 1999-04-20

; PRIOR APPLICATION NUMBER: US 09/314,464

; PRIOR FILING DATE: 1999-05-18

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15

; LENGTH: 2740

; TYPE: DNA

; ORGANISM: Human

US-09-594-669-15

Query Match

Best Local Similarity 14.6%; Score 449.8; DB 3; Length 2740;

Matches 744; Conservative 62.4%; Pred. No. 1.5e-105;

Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;

QY 1300 TGAAGAGCTGGAGCTGACCTGACGAGCAATCGGAGAGCGACGCGCCGACAGCGGAG 1359

Db |||||

638 TGAAGAGCTGGAGCAATGAGAAACAGCGAGAGAGAGAGGCGCCGAACTCTGAAA 697

QY 1360 TGAAGAGAGAGAGGTGGCGCTGATGAACAGGATCCGGGCAATCCAACTGGGAGCG 1419

Db |||||

698 TGAGATGAAGAG-----CTCAGGAGTATGACAGTAGTTTTTCCAACTGGAGTTG 751

QY 1420 CGCAATATGATACGCAATATCAGAGCAGCTGGAATTTGTGCGGTGCTCGATGGCCAGG 1479

Db |||||

752 CCGCAATGATTAAGAAATTCGGGCTACTTTGGAATGTCATCCACTTACTATGACTGATC 811

QY 1480 CCGTCGATCACCATCAGATCAGTGTGCGGCAAGGCTCCCATAGCCGCAAGGAGG 1539

Db |||||

812 CTATCGAAGAGCACAGAAATATGTGTGTGTTAGGAAACGCCCACTGAAATAAGCAAGAT 871

QY 1540 TCAATCGAAGGAGATCGATGTCATTTCCGTCGCGCAAGGACATGCTCATCGTGACG 1599

Db |||||

872 TGGCCAGAGAGAAATGATGATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 931

QY 1600 AGCGCGCAGCAAGGTGAGCTCACCAGTTCCTGGAGAACCAAGTTCGCTTCGACT 1659  
Db |||||  
932 AACCCAAAGTTGAAGTGGACTTAACAAAGTATCTGGAGAACCAAGCATTTCTGCTTGA 991  
QY 1660 ACSCCTTCAACGACAGCTGGGCAAAATGCCATATGTTATACAAATACAGCCAGCCGTTGG 1719  
Db |||||  
992 TTGCATTTGATGAAACAGCTTTCGAATGAAGTTGTCTACAGGTTTACAGCAGGCGCACTGG 1051  
QY 1720 TGAACCACTTTTCGAGGGCGGAATGGCGACGTCTTCGCTACGGCCAGACGGGATCGG 1779  
Db |||||  
1052 TACAGACATCTTTTGAAGTGGAAACAACTTTGTTTGCATATGGCCAGACAGGAAGTG 1111  
QY 1780 GCAAAACGACACATCGGCGGTGAGTTTAAATGGAAGGTGCAAGACTGCAAGAACCGCA 1839  
Db |||||  
1112 GCAAGACACATATATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGGA 1171  
QY 1840 TCTACGCCATGGCGGCAAGGATGTTTGTGACCTGATATATGCGCGTTACCGCGCCA 1899  
Db |||||  
1172 TCTATGCCATGGCGCTCCCGGAGCTTCTCTCTGGAAGAAATCAACCTTGTACCGGAAGT 1231  
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTCCTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC 1959  
Db |||||  
1232 TGGGCTTGAAGTCTATGTGACATCTTTCAGATCTTACATGGGAAGCTGTTTGAACCTGC 1291  
QY 1960 TGTCCGACAAAGCAGAACTGCGGCTCTGGAGATGTTAAACAGCAAGTSCAGGTGTTGG 2019  
Db |||||  
1292 TCAACAAAGAAAGCCAAAGCTGCGGCTGCGAGGAGCGCAAGCAAGGTGCAAGTGGTGG 1351  
QY 2020 GACTACCGAGAAAGGTGGTTCGATGGCGTGGAGAGTACTGAAGCTCATCCAGCAGCGCA 2079  
Db |||||  
1352 GGTGCGAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGCTCGACATGGGCA 1411  
QY 2080 ATGCTGCCCGCAACATCCGCGCAGAGCTGCGCAACTCCAAATTCGTCGCTGCGGTCG 2139  
Db |||||  
1412 GCGCTGCGAGAACCTCTGGGAGACATTTGCCAACTCCAAATTCCTCCGCTCCACCGCT 1471  
QY 2140 TTTTCCAGATTTGCTGCGCGCCGCGCTCGACGAGATGCCATGGCAAGTTCCTCGTTCA 2199  
Db |||||  
1472 GCTTCCAAATTTTCTTCGAGCTAAAG-----GAGATGCATGCGCAAGTTCCTTTGG 1525  
QY 2200 TCGATCTGGCGGCAATGAGCGGGGGTGGACACTTCTCGGCGGATCGGCGAGCGGTA 2259  
Db |||||  
1526 TAGTCTGGCAGGAATGAGCGGCGCAGACACTTCCAGTGTCTGACCGGCGAGCCGCA 1585  
QY 2260 TGGAGGTGCGCAGATTAACAATCGCTGCTGCGCCCTCAAGGAGTGCATTCGTCGTTGG 2319  
Db |||||  
1586 TGGAGGCGCAGAAATCAACAAGATCTTTAGCCCTGAGGAGTGCATCAGGGCCCTGG 1645  
QY 2320 GCAACAGTCCGCCCACTTTGCCCTTCCGTCTCTCCAAACTCACCAGGTGCTGCGGACT 2379  
Db |||||  
1646 GACAGAAACAAGGCTCACACCCCGTTCCGTCGAGGCAAGCTGACACAGGTGCTGAGGACT 1705  
QY 2380 CTTTCAATTTGGCAGAGCAAGACGTCGATATAGCCATGATCTCGCGGAGCTTAGCT 2439  
Db |||||  
1706 CTTTCAATTTGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCAGGCATAGCT 1765  
QY 2440 CTGCGAGCACAGCTCAACAGCTGCGTATGCGCATGCTGTCAAGGAGCTG 2492  
Db |||||  
1766 CTTGGAATATATCTTTAAACACCCCTGAGATATGACAGAGGGTCAAGGAGCTG 1818

## RESULT 12

US-09-595-684B-32

; Sequence 32, Application US/09595684B

; Patent No. 6544766

; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe

; APPLICANT: Ohashi, Cara

; APPLICANT: Sakowicz, Roman

; APPLICANT: Vaisberg, Eugeni

; APPLICANT: Wood, Kenneth

; APPLICANT: Yu, Ming

; TITLE OF INVENTION: Human kinesins and methods of producing

```

; TITLE OF INVENTION: and purifying human kinases
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 2740
; TYPE: DNA
; ORGANISM: Human
US-09-595-684B-32

```

Query Match	14.6%;	Score 449.8;	DB 3;	Length 2740;
Best Local Similarity	62.4%;	Prod. No. 1.5e-105;		
Matches 744;	Conservative 0;	Mismatches 437;	Indels 12;	Gaps 2;
QY	1300	TGAAGAGGTGGAGCGACTGAAGGAGAAATCGCGAGAGCGACGCGCGGACAGCGCCGAGA	1359	
DB	638	TGAAGGNAGTGGAAANAATGAGANAACAAGCGAGAGAGAGAGAGGCCCCAGAACTCTGAAA	697	
QY	1360	TGAAGGAGGAGAAGGTGGCGCTGATGAACACAGGATCCGGGCAATCCAAACTGGGAGACGG	1419	
DB	698	TGAGAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTCTTTTCCAAACTGGGAATTG	751	
QY	1420	CGCAATGATACCGGATATCAGAGCAGCTGGAATTTGTGCCGCTGCTCGATGGCCAGG	1479	
DB	752	CCGCAATGATTAAGAANAATTCGGGCTACTTTGGAATGTCACTTACTTATGACTGATC	811	
QY	1480	CCGTGATGACCATCAGATCACAGTGTGCGTGGCGAAGCGCTCCCACTAGCGCAAGGAGG	1539	
DB	812	CTATCGAAGAGCACAGAATATGTCTGTCTAGGAAACGCCCACTGAATAAGCAAGAAT	871	
QY	1540	TCAATCGCAAGGAGATCGATGTCAATTCGGTGGCGCGCAAGGACATGCTCATCGTGCAGC	1599	
DB	872	TGCCCAAGAAGAANAATGATGTGAATTCCAATCTCTACCAAGTGTCTCTCTTGGTACATG	931	
QY	1600	AGCGCGCGACCAAGGTCGACCTCACCAAGTTCTCTGGAGAACCAAGTTTTCGCTTCGACT	1659	
DB	932	AACCCAAGTTGAAGTGGACTTAAACAAGTATCTGGAGAACCAAGCAATTCGCTTTGACT	991	
QY	1660	AGCCTTCAACGACACGTGCGCACAATGCCATGGTATACAAATACAGCCAAAGCCGTTGG	1719	
DB	992	TTGCAATTTGATGAACAGACTTCGAATGAAGTTGTCTACAGGTTTACAGCAAGGCCACTGG	1051	
QY	1720	TGAAAACCATTTTCGAGGCGGGAATGCGCAGCTGCTTCGCTACGGCCAGCGGATCGG	1779	
DB	1052	TACAGACATCTTTGAAGTGGAAAGCAACTTGTCTTGCATATGGCCGACAGGAAGTG	1111	
QY	1780	GCAAAACGCACACCATGGGCGGTGAGTTTAAATGGAAGGTGCAAGGATGCAAGAACCGCA	1839	
DB	1112	GCAAGACACATACTATGGGCGGAGACTCTCTCGGAAAGGCCAGAAATGCATCCAAAGGA	1171	
QY	1840	TCTACGCCATGGGCGGCAAGGATGTCTTTGTGACCTGAAATATGCCGCGTTACCGCGCCA	1899	
DB	1172	TCTATGCCATGGGCTCCCGGAGCTTCTCTCTGAAGAAATCAACCTGCTACCGGAAGT	1231	
QY	1900	TGAATCTAGTCTCGCTCGGCGAGTTCTTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC	1959	
DB	1232	TGGGCTGGAAGTCTATGTGACATCTTCGAGATCTCAATTGGGAAGCTGTTTGACTGC	1291	
QY	1960	TGTCCGACAAAGCAGAAATCGCGGTCTGGAGAGTGGTAAACAGCAAGTGCAGGTGGTG	2019	
DB	1292	TCAACAAGAAGGCCAAGCTGCGCGTCTGGAGAGCGCAAGCAACAGGTGCAAGTGGTG	1351	
QY	2020	GACTCACCCAGAGGTGGTTCGATGGCGTGCAGGAGGTACTGAGGCTCATCCAGCAGCGCA	2079	
DB	1352	GGCTGCAGAGGCATCTGGTTAACTCTGTGATGATGTCAATGATGCTCGAATGGGCA	1411	
QY	2080	ATGTGCGCCGAACATCCGCGCCAGAGCTCGGCCAACTCCAAATTCGTGCGGTTGCACGCG	2139	
DB	1412	GGCGCTGCAGAACTCTGGGCGACATTTGGCAACTCCAAATCTCCGCTCCCAAGCT	1471	

```

RESULT 13
US-10-112-432-15
; Sequence 15, Application US/10112432
; Patent No. 6638754
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6638754e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/112.432
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2740
; TYPE: DNA
; ORGANISM: Human
US-10-112-432-15

```

Query Match	14.6%	Score 449.8	DB 3	Length 2740
Best Local Similarity	62.4%	Pred. No. 1.5e-105		
Matches 744	Conservative 0	Mismatches 437	Indels 12	Gaps 2
Qy	1300	TGAAGAGGTGGAGCGACTCAAGGAGAATCGCGAGAAGCGACGCGCCGACAGGCCGAGA	1359	
Db	638	TGAAGGAGTGGAAAAATGAAGAACAGCGCAGAGAGAGAGAGGCCGAGAACTCTGAAA	697	
Qy	1360	TGAAGGAGGAGAAGGTGGCGCTGATGAACAGGATCCGGGCAATCCAAACTGGGAGACGG	1419	
Db	698	TGAGAATGAAGAGAG-----CTCAGAGTATCAGCAGTAGTTTTTCCAAACTGGGAATTTG	751	
Qy	1420	CGCAATGATACCGGAATATCAGAGCAGCGCTGGAATTTGTGCGCGTCTCGATGGCCAGG	1479	
Db	752	CCGCAATGATTAAGAAATTTCCGGCTACTTTGGAAATGTCAATCCACTTACTATGACTGATC	811	
Qy	1480	CCGTCGATGACCATCAGATCCAGTGTGGCTGGCGCAAGCGTCCCATTTAGCCGAAGGAGG	1539	
Db	812	CTATCGAAGAGCAGAAATATGTCTGTGTATAGGAAACGCCCACTGAATTAAGCAGAAGAT	871	



```
QY 2140 TTTTCAGATTGTGTCGGCCGACAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTCA 2199
Db 1472 GCTTCCAAATATTCTTCGAGCTAAAG-----GAGAATGATGGCAAGTTCTCTTGG 1525
QY 2200 TCGATCTGGCGGCAATGAGCGGGCGTGGACACTTCTCGGCCGATCGCAGACGGTGA 2259
Db 1526 TAGATCTGGCAGGGAATGAGCGGCGCAGACACTTCCAGTGTGACCGCGCAGACCGCA 1585
QY 2260 TGGAGGGTCCGAGATTAAACAAATCGCTGCTGGCCCTCAAGAGTGCATTCTGCGTTGG 2319
Db 1586 TGGAGGGCGCAGAATCAACAAGAGTCTCTTAGCCCTGAAGGAGTGCATCAGGGGCCCTGG 1645
QY 2320 GCAACAGTCGCGCCACTTGCCTTCCCGTGTCTCAAACTCACCAGGTGCTGCGCGACT 2379
Db 1646 GACAGAAACAAGCTCACACCCGTTCCGTGAGAGCAGCTGACAGGTGCTGAGGGACT 1705
QY 2380 CTTTCATTGGCGAGAGAGCAGAGTGCATGATAGCCATGATCTGCGCGGACTTAGCT 2439
Db 1706 CTTTCATTGGGGAACCTTAGGACTTGCATGATTGCCACGATCTCACCAGGCATAAGCT 1765
QY 2440 CTTGGAGCACACGCTCAACACGCTGCGCTATGCGGATCGTGCAGGAGCTG 2492
Db 1766 CCTGTGAATATCTTTAAACCCCTGAGATATGCGACAGGGGTCAAGGAGCTG 1818
```

```
RESULT 15
US-09-594-669-5
; Sequence 5, Application US/09594669
; Patent No. 6331424
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6331424e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1042
; CURRENT APPLICATION NUMBER: US/09/594,669
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: Human
US-09-594-669-5
```

```
Query Match 13.8%; Score 425.6; DB 3; Length 1088;
Best Local Similarity 65.2%; Pred. No. 1.7e-99;
Matches 645; Conservative 0; Mismatches 339; Indels 6; Gaps 1;

QY 1503 GTGTGCGTCGCGCAAGCGTCCCATTAGCCGCAAGGAGTCAATCGCAAGGAGATCGATGTC 1562
Db 15 GTCTGTGTTAGGAACGCCACTGAATAGCAAGATTTGGCCAAAGAAATTTGATGTG 74
QY 1563 ATTTGGGTGCGCGCAAGGACATGCTCATCTGTGCACGAGCGCGCAGCAAGGTGACCTC 1622
Db 75 ATTTCCATTCCTAGCAAGTGTCTCCTCTTGGTACATGAACCCCAAGTTGAAAGTGSACTTA 134
QY 1623 ACCAAGTTCCTGGAGAACCAAGTTTCCTTCGACTACGCTTCAACGACACGTCGGAC 1682
Db 135 ACAAAGTATCTGGAGAACCAAGCATTTCTGCTTTGACTTTGCAITTTGATGAACACGCTTCG 194
QY 1683 AATGCCATGTGTATACAAATACACAGCCGATTTGGTGAATAACATTTTCGAGGGCGGA 1742
Db 195 AATGAAGTTGTCTACAGTTTACAGCAAGGCCACTGGTACAGACAACTTTTGAAGTTGGA 254
QY 1743 ATGGCGAGTGTCTTCGCTAGCGGCAGACGCGGATCGGGGAAAAACGACACCATGGGCGGT 1802
Db 255 AAAGCAACTTGTGTTTGCATATGGCCAGACAGGAAGTGGCAAGACACATATCTATGGGCGGA 314
```

```
QY 1803 GAGTTTAAATGGAAGGTGCAGACTGCAAGACGCATCTACGCCATCGCGGCCAAGGAT 1862
Db 315 GACCTCTCTGGGAAAGCCAGAAATCATCCAAAGGGATCTATGCCATGGCTTCCGGGAC 374
QY 1863 GTCTTTGTGACCCCTGAATATGTCGCGGTTTACCGCGCCATGAATCTAGTCTGTCTCGGCCAGT 1922
Db 375 GTCTTCTCTGAAGAATCAACCCCTGTCTACCGGAAGTTGGGCTTGGAGTCTATGTGACA 434
QY 1923 TTTTGTGAGATTACAGTGGCAAGGTCTTCGATCTTCTGTCCGACAAAGCAAGAACTGCGC 1982
Db 435 TTTTTCGAGATCTACAATGGGAAGCTGTTTGAACCTGCTCAACAAGAGGCAAGCTGCGC 494
QY 1983 GTCTTGGAGGATGGTAAACAGCAAGTGCAGTGGTGGGACTCACCGAAGAGGTGTCAT 2042
Db 495 GTCTTGGAGGATGGCAAGCAAGTGCAGTGGTGGGCTGCAAGGAGCATCTCGTTTAAAC 554
QY 2043 GCGGTCGAGGAGGTACTGAAGCTATCCAGCAAGCAATGCTGCCGCAACATCCGCGCAG 2102
Db 555 TCTGCTGATGATGATCATCAAGATGATCGACATGGCAGCGCTGCAGAACCTCTGGGCGAG 614
QY 2103 AGCTCGGCCAACTCCAAATTCGTGCGGTTCGCAAGCCGTTTTCAGATTTGTGTCGCGCG 2162
Db 615 ACATTTGCCAACTCCAATTTCTCCGCTCCCAAGCGTGTCTTCCAAATTTATTTCTTGAGCT 674
QY 2163 CAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTTCATCGATCTGGCGGGCAATGAGCGG 2222
Db 675 AAAGG-----GAGAAATGCTGGCAAGTTCTTTTGGTAGATCTGGCAGGGAATGAGCGA 728
QY 2223 GCGGTGGACACTTCTCTCGGCCGATCGGCAGACGCGTATGGAGGGTCCGAGATTTAAACAA 2282
Db 729 GCGCGGACACTTCCAGTGTCTGACCGGCAGACCCGCGATGGAGGGCGCAGAAATCAACAAG 788
QY 2283 TCGCTGCTGGCCCTCAAGGAGTGCATTGTCGTGGTGGGCAACAGTGGCCCACTTTGCC 2342
Db 789 AGTCTCTTAGCCCTGAAGGAGTGCATCAGGCGCCCTGGGACAGAACAGGCTCACACCCCG 848
QY 2343 TTTCCGTGTCTCCAACTCACCCAGGTGTCGCGCACTCGTTTCATTTGGCGAGAGGCAAG 2402
Db 849 TTCCGTGAGAGCAAGCTGCACAGGTGCTGAGGGAATCTCTTTCATTTGGGGAGAACTCTAGG 908
QY 2403 ACGTGCATGATAGCATGATCTCGCCGGGACTTAGCTCTCGGAGCACACGCTCAACACG 2462
Db 909 ACTTGCATGATTGCCAGATCTCACAGGAGCAATAGCTCCTGTGATATATCTTTAAACACC 968
QY 2463 CTGCGCTATGCGGATCGTTCAGGAGCTG 2492
Db 969 CTGAGATATGCAGACAGGGTCAAGGAGCTG 998
```

Search completed: May 12, 2006, 21:18:16  
Job time : 532 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 17:35:46 ; Search time 11790 Seconds  
(without alignments)  
12226.548 Million cell updates/sec

Title: US-10-840-060-142

Perfect score: 3081

Sequence: 1 aaactaaaaattgtgttc.....gccgaatggcaacgcttagt 3081

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_est3.\*

4: gb\_hc.\*

5: gb\_est4.\*

6: gb\_est5.\*

7: gb\_est6.\*

8: gb\_est7.\*

9: gb\_gss1.\*

10: gb\_gss2.\*

11: gb\_gss3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	917	29.8	1101	10	CNS0011P
2	761.6	24.7	779	2	BP499568
3	752.6	24.4	759	2	BP484448
4	750.2	24.3	763	2	BP498461
5	720.4	23.4	722	2	BP489668
6	718	23.3	1105	8	DR122830
7	692	22.5	1153	8	DR135726
8	685.8	22.3	720	2	BP498964
9	683.6	22.2	702	2	BG633810
10	680.6	22.1	687	2	BP495560
11	678	22.0	686	2	BP499204
12	674.4	21.9	689	3	BI635599
13	663.8	21.5	679	2	BP495237
14	663.6	21.5	670	2	BG634444
15	662.4	21.5	664	3	BI609735
16	659.8	21.4	744	2	BF505411
17	659	21.4	660	3	BI592847
18	656	21.3	656	1	AI259349
19	641	20.8	641	2	BP488088
20	629	20.4	629	1	AI388100
21	625.2	20.3	630	2	BF503725
22	619.6	20.1	626	2	BP491708

23	605.8	19.7	719	1	AI546445
24	599	19.4	651	3	BI372858
25	591.6	19.2	598	2	BF506249
26	590.4	19.2	695	1	AA949737
27	589.2	19.1	605	3	BI354473
28	589	19.1	740	1	AA952135
29	588.4	19.1	682	2	BF488658
30	587.8	19.1	591	2	BF492432
31	582.6	18.9	749	2	BF506266
32	582.2	18.9	642	1	AI546390
33	577	18.7	586	1	AI532286
34	575.2	18.7	721	7	CK660542
35	575	18.7	575	1	AI515654
36	574	18.6	574	3	BI609578
37	574	18.6	574	3	BI609874
38	574	18.6	575	3	BI609406
39	573.2	18.6	587	1	AI533352
40	571.8	18.6	587	2	BP492824
C 41	570	18.5	664	10	AG924280
42	568	18.4	568	2	BF491744
43	563.4	18.3	565	3	BI609731
44	560.8	18.2	572	3	BI610275
45	558.4	18.1	597	2	BE975372

#### ALIGNMENTS

RESULT 1

CNS0011P 1101 bp DNA linear GSS 03-JUN-1999

LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #

DEFINITION BACR04G01 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL061056

VERSION AL061056.1 GI:4939519

KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT - Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org/TheBDGP/Drosophila> melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y; cn bw sp. The same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

#### FEATURES

Location/Qualifiers

1..1101

/organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone="BACR04G01"

/clone\_lib="RPCI-98"

/note="end : TET3"

#### ORIGIN

Query Match		29.8%; Score 917; DB 10; Length 1101;
Best Local Similarity		90.9%; Pred. No. 4e-227;
Matches		957; Conservative 41; Mismatches 52; Indels 3; Gaps 3;
QY	269	CACACACACACACTGGGACACTCGAAAAAACAACAAAAACAACAGACGCTGTGCGGACCTCGA 328
DB	5	CACACACACACTTAGAACACATCAAAAAAACAACAAAAACAACAGCTGTGCGGACCTCAA 64
QY	329	CAAGCAGTGTGGTTCGATTGTTGTCATATGCTTGAATTCGGTTCGGGCTTGTGTTTC 388
DB	65	CAACATTTTYYKTYKTCATTTATTTTTCATGCTTAAATTCGTTTCGGGCTTATTTTC 124
QY	389	ACAAGTTTATCGCTCGTCAAGAAACAACAGAAATAAAATATTTTCGACCTAAAAATCTG 448
DB	125	ACAATTTTATCGCTCTTCAAAAAACAACAAATAAAATATTTTCGACCTAAAAATCTA 184
QY	449	ACTAAATGTGTTTTTGTATTGTTATTTATTTATTTAGGCACATTTTTCACACACACAGTAG 508
DB	185	ACTAAATTTTTTTTTTTTTTTATTTATTTATTTATTTAGGCACATTTTTCACACACACAGTA - 243
QY	509	TTACTACATCTACGACTAACGGAACCTCTCTCGAAGCAGTGAAGTTGCTGTCCATCAA 568
DB	244	TTACTACATCTACGACTAACGGAACCTCTCTCGAAGCAGTGAAGTTGCTGTCCATCAA 303
QY	569	GCAGTACTCGGAGTTAAACGACGAGTAAAGCCGGAGAAAGAGAGATCGGTGGAGAAT 628
DB	304	GCAGTACTCGGAGTTAAACGACGAGTAAAGCCGGAGAAAGAGAGATCGGTGGAGAAT 363
QY	629	AGAGATATACAGTGGAGTCAAGAGAAAGGATCATGGAATGATTACGTTGGGGCAGAG 688
DB	364	AGAGATATACAGTGGAGTCAAGAGAAAGGATCATGGAATGATTACGTTGGGGCAGAG 423
QY	689	CGTCAAGATCAAGCGGACGATGCGCGCTCCATGCGCGTGGTGGGATCAACCA 748
DB	424	CGTCAAGATCAAGCGGACGATGCGCGCTCCATGCGCGTGGTGGGATCAACCA 483
QY	749	GTCGGGCAAGTGCATCACAGTTCGAATGTTACGAGCGCGGCAACGAAGCGCAAGGAGT 808
DB	484	GTCGGGCAAGTGCATCACAGTTCGAATGTTACGAGCGCGGCAACGAAGCGCAAGGAGT 543
QY	809	AGAACTGGACGCATCTACGCTCAATTCGGAGCTTAATCAAGATGACTGTGCAACAGCA 868
DB	544	AGAACTGGACGCATCTACGCTCAATTCGGAGCTTAATCAAGATGACTGTGCAACAGCA 603
QY	869	CGCGCCCGCGAGCCCAAGAAACAAGCCACCGCGCGATCAACTCTCGGTTAATCCAC 928
DB	604	CGCGCCCGCGAGCCCAAGAAACAAGCCACCGCGCGATCAACTCTCGGTTAATCCAC 663
QY	929	ACAATCGGCTATCGGTGGCAATCTCACAGCCGCTATGACCATGCGCGGAAACATGCTGAA 988
DB	664	ACAATCGGCTATCGGTGGCAATCTCACAGCCGCTATGACCATGCGCGGAAACATGCTGAA 723
QY	989	CAAGATCCAGAAAGCCAGTTCGATTCCTCAATTCGATTGTTCAGCAGCAATAGCTGAATC 1048
DB	724	CAAGATCCAGAAAGCCAGTTCGATTCCTCAATTCGATTGTTCAGCAGCAATAGCTGAATC 783
QY	1049	AAACAGCACTCCCAACATACGCGCGCGAGTGTGGCAACCAACACACGTCGACGCCAC 1108
DB	784	AAAMAGCACTCCCAACATACGCGCGCGAGTGTGGCA-CWCAACSTFCGACGCCAC 842
QY	1109	TGGATTACAGCGTCCAGGTAATCTCGAAGCTGTACCGGCGCAGCAGACAGCAAGGATCGC 1168
DB	843	TGGATTACAGCGTCCAGGTAATCTCGAAGCTGTACCGGCGCAGCAGACAGCAAGGATCGC 902
QY	1169	CTCGGCGGTGCTTAATAACATTTGCCCCAATCCAGCGCGGACCGCAGTGTGCTCGGCG 1228
DB	903	CYCGCGGTGCTTAATAACATTTGCCCCAATCCAGCGCGGACCGCAGTGTGCTCGGCG 962
QY	1229	GGCACAAGGATCGCCATCTCGGCCCAACCCAGGGAGCTGGCGGCTAGTACCCGGCG 1288
DB	963	SGCACAARGAGTGR-CMWYKSGGCCAAWACCMRGAGCTTTCGSGCSTATWACCCGCSA 1021
QY	1289	ATCGCACGCAATTGAAAGAGGTGGAGCGACTGAA 1321

DB	1022	TYCMCCATTWAAAGGTGARCSTACTGAAGAA 1054
RESULT 2		
BF499568		779 bp mRNA linear EST 19-APR-2001
LOCUS		AT14153.Sprime AT Drosophila melanogaster adult testes pOTB7
DEFINITION		Drosophila melanogaster cDNA clone AT14153 5 similar to CGI453:
		PB00001453 'motor protein' located on: X 10A6-10A7;: 04/09/2001,
		mRNA sequence.
ACCESSION		BF499568
VERSION		BF499568.2 GI:113691428
KEYWORDS		EST.
SOURCE		Drosophila melanogaster (fruit fly)
ORGANISM		Drosophila melanogaster
		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
		Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
		Ephydroidea; Drosophilidae; Drosophila.
REFERENCE		1 (bases 1 to 779)
AUTHORS		Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
		Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
		Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
		Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C. J., Nunoo, J.,
		Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
		Lewis, S. E., Celniker, S. and Rubin, G. M.
TITLE		BDGP/HMI AT Drosophila EST Project
JOURNAL		Unpublished (2000)
COMMENT		On Dec 6, 2000 this sequence version replaced gi:11582869.
		Contact: Stapleton, M.
		BDGP
		Lawrence Berkeley National Lab
		One Cyclotron Rd, Berkeley, CA 94720, USA
		Fax: 510 486 6798
		Email: <a href="http://www.fruitfly.org/EST">http://www.fruitfly.org/EST</a> , <a href="mailto:est@fruitfly.berkeley.edu">est@fruitfly.berkeley.edu</a>
		hit genomic AB003485; arm:X [10878159,11180508]
		estimated-cyto:10A3-10B11: 04/09/2001
		Plate: AT141 row: E column: 5
		High quality sequence stop: 689.
FEATURES		Location/Qualifiers
		1..779
		/organism="Drosophila melanogaster"
		/mol_type="mRNA"
		/db_xref="taxon:7227"
		/clone="AT14153"
		/sex="male"
		/dev_stage="0-3 day old Ore-R males"
		/lab_host="Plates AT.10-AT.120: DHS-alpha. Plates
		AT.121-AT.319: DHS-alpha Tona"
		/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
		/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcorI;
		Site: 2: XhoI; The mRNA for the testis library was made
		from testes and seminal vesicles hand dissected from 0-3
		day old Ore-R males. RNA kindly provided by the lab of
		Margaret Fuller. Sized fractionated cDNAs were directly
		ligated into pOTB7. Plasmid cDNA library."
ORIGIN		
		Query Match 24.7%; Score 761.6; DB 2; Length 779;
		Best Local Similarity 99.4%; Pred. No. 1.1e-186;
		Matches 775; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY	387	CCACAGTTTATCGTTCGTCAGAAACACAGAAATAAAATTTTTCGACCTAAAAATC 446
DB	1	CCACAGTTTATCGTTCGTCAGAAACACAGAAATAAAATTTTTCGACCTAAAAATC 60
QY	447	TGACTAAATTTGTTTTTTTGTATTATTTAGGCACATTTTTCACACCAACCGT 506
DB	61	TGACTAAATTTGTTTTTTTGTATTATTTAGGCACATTTTTCACACCAACCGT 120
QY	507	AGTTACTACATCTACGACTAACCGAACTCCTCTCTCAAGCAGTGGAAAGTTGTCATC 566
DB	121	AGTTACTACATCTACGACTAACCGAACTCCTCTCTCAAGCAGTGGAAAGTTGTCATC 180



QY 567 AAGCAGTACTCGGAGTTAACCGCAGGATAAGCGGAGAAAGAGAGATCGGTGGAGA 626  
 Db |||||  
 QY 181 AAGCAGTATACCGGAGTTAACCGCAGGATAAGCGGAGAAAGAGAGATCGGTGGAGA 240  
 Db |||||  
 QY 627 ATAGAGATATACAGGTGAGTCAAGAGAGAAAGGATCATGGACATGATTTACGGTGGGCGAG 686  
 Db |||||  
 QY 241 ATAGAGATATACAGGTGAGTCAAGAGAGAAAGGATCATGGACATGATTTACGGTGGGCGAG 300  
 Db |||||  
 QY 687 AGCGTCAAGATCAAGCGGACCGATGGCGCGTCCACATGGCGGTGGTGGCGGTGATCAAC 746  
 Db |||||  
 QY 301 AGCGTCAAGATCAAGCGGACCGATGGCGCGTCCACATGGCGGTGGTGGCGGTGATCAAC 360  
 Db |||||  
 QY 747 CAGTGGGCGAGTGCATCAGATCGATGGTACGAGCGCGCGGAAACGAAGGCGAAGGAG 806  
 Db |||||  
 QY 361 CAGTGGGCGAAGTGCATCAGATCGATGGTACGAGCGCGCGGAAACGAAGGCGAAGGAG 420  
 Db |||||  
 QY 807 GTAGAACTGGAGCGCATACTCAGCTCAATCCGAGGCTAATGCAAGATACTGTGCAACAG 866  
 Db |||||  
 QY 421 GTAGAACTGGAGCGCATACTCAGCTCAATCCGAGGCTAATGCAAGATACTGTGCAACAG 480  
 Db |||||  
 QY 867 CAGCGCGCGCGGAGCGCAAGAAAGCCACCGCGCGATGAACCTCTCGCGTAATCCC 926  
 Db |||||  
 QY 481 CAGCGCGCGCGGAGCGCAAGAAAGCCACCGCGCGATGAACCTCTCGCGTAATCCC 540  
 Db |||||  
 QY 927 ACACAATCGGTATCGGTGGCAATCTACAGCGCGTATGACCATGGCGCGGAAACATCGTG 986  
 Db |||||  
 QY 541 ACACAATCGGTATCGGTGGCAATCTACAGCGCGTATGACCATGGCGCGGAAACATCGTG 600  
 Db |||||  
 QY 987 AACAAAGATCCAGGAAGCCAGTGCATTCCTCAATCCGATTTGTCAGCAGCAATAGCGTGAAT 1046  
 Db |||||  
 QY 601 AACAAATCCAGGAAGCCAGTGCATTCCTCAATCCGATTTGTCAGCAGCAATAGCGTGAAT 660  
 Db |||||  
 QY 1047 ACAACAGCACTCCAACTACCGCGCGGAGGTGTGGCAGCAACAAAGTCCAGCAGCC 1106  
 Db |||||  
 QY 661 ACAACAGCACTCCAACTACCGCGCGGAGGTGTGGCAGCAACAAAGTCCAGCAGCC 719  
 Db |||||  
 QY 1107 ACTGGATTACGGTCCAGCGTACTCGCAAGCTCTACCGGCGGAGCGAGCAAGGATC 1166  
 Db |||||  
 QY 720 ACTGGATTACGGTCCAGCGTACTCGCAAGCTCTACCGGCGGAGCGAGCAAGGATC 779  
 Db |||||

## RESULT 3

Bf486448 759 bp mRNA linear EST 19-APR-2001  
 LOCUS AT20744.5prime AT Drosophila melanogaster adult testes pOTB7  
 DEFINITION Drosophila melanogaster cDNA clone AT20744 5 similar to CG1453;  
 Fnan0001453 'motor protein' located on: X 10A6-10A7; 04/09/2001,  
 mRNA sequence.

## ACCESSION

Bf486448

Bf486448.2 GI:13698328

EST.

## VERSION

Drosophila melanogaster

## KEYWORDS

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

## SOURCE

1 (bases 1 to 759)

## ORGANISM

Drosophila melanogaster (fruit fly)

## REFERENCE

Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,

Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,

L., P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,

Paclet, J., Paragav, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,

Lewis, S.E., Celniker, S. and Rubin, G.M.

BDGP/HMI AT Drosophila EST Project

## AUTHORS

Unpublished (2000)

On Dec 6, 2000 this sequence version replaced gi:11569749.

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)

hit genomic AB003485: arm:X [10878159,11180508]  
 estimated-cyto:10A3-10B11: 04/09/2001  
 Plate: AT.207 row: D column: 8  
 High quality sequence stop: 734.

## FEATURES

## source

1..759

/organism="Drosophila melanogaster"

/mol\_type="mRNA"

/db\_xref="taxon:7227"

/clone="AT20744"

/sex="male"

/dev\_stage="0-3 day old Ore-R males"

/lab\_host="Plates AT.10-AT.120: DH5-alpha. Plates

AT.121-AT.319: DH5-alpha Tona"

/clone\_lib="Ar Drosophila melanogaster adult testes pOTB7"

/notes="Organ: ADULT testes; Vector: pOTB7; Site 1: EcoRI;

Site 2: XhoI; The mRNA for the testis library was made

from testes and seminal vesicles hand dissected from 0-3

day old Ore-R males. RNA kindly provided by the lab of

Margaret Fuller. Sized fractionated cDNAs were directly

ligated into pOTB7. Plasmid cDNA library."

## ORIGIN

Query Match

Best Local Similarity

Matches

755; Conservative

0; Mismatches

4; Indels

0; Gaps

0;

QY

409

GAAACACGAAATAAAATATTTTCGACCTAAATAAAATCTGACTAAATTTGTTTTTTTGT 468

Db

1

GAAACACGAAATAAAATATTTTCGACCTAAATAAAATCTGACTAAATTTGTTTTTTTGT 60

QY

469

TATGTATTTATTTAGGCACATTTTGCACACCAACCAAGCTAGTTACTACATCTACGACTAAC 528

Db

61

TATGTATTTATTTAGGCACATTTTGCACACCAACCAAGCTAGTTACTACATCTACGACTAAC 120

QY

529

GGAACCTCTCTCGACAGCAGTGAAGTTGCTTCCATCAAGCAGTACTCGAGTTAAGCG 588

Db

121

GGAACCTCTCTCGACAGCAGTGAAGTTGCTTCCATCAAGCAGTACTCGAGTTAAGCG 180

QY

589

AGGATAAGCGCGGAGAAAGAGAAAGAGATCGGTGGAGAAATAGAGATATACAGTTGGAGTC 648

Db

181

AGGATAAGCGCGGAGAAAGAGAAAGAGATCGGTGGAGAAATAGAGATATACAGTTGGAGTC 240

QY

649

AAAGAGAGAGGATCATGACATGATTTACGTTGGGCGAGAGCGTCAAGATCAAGCGGACGG 708

Db

241

AAAGAGAGAGGATCATGACATGATTTACGTTGGGCGAGAGCGTCAAGATCAAGCGGACGG 300

QY

709

ATGCGCGGTCCACATGCGCGGTGATCAACCAAGTGGGCAAGTGCATCAAG 768

Db

301

ATGCGCGGTCCACATGCGCGGTGATCAACCAAGTGGGCAAGTGCATCAAG 360

QY

769

TCGAATGGTACGAGCGCGGAAACGAAGGCGAGAGGTAGAACTCGACGCCATCTCA 828

Db

361

TCGAATGGTACGAGCGCGGAAACGAAGGCGAGAGGTAGAACTCGACGCCATCTCA 420

QY

829

CGCTCAATCCGAGCTAATGCAAGATCTGTGAAACAGCAGCGCGCGCGGAGCCCAAGA 888

Db

421

CGCTCAATCCGAGCTAATGCAAGATCTGTGAAACAGCAGCGCGCGCGGAGCCCAAGA 480

QY

889

AACAGCGACCGCGCGGATGAACTCTCGGTTAATCCACACAATCGGCTATCGGTGGCA 948

Db

481

AACAGCGACCGCGCGGATGAACTCTCGGTTAATCCACACAATCGGCTATCGGTGGCA 540

QY

949

ATCTCACCAGCGGTATGACCATGCGCGGAAACATGCTGAACAGATCCAGGAAGCCAGT 1008

Db

541

ATCTCACCAGCGGTATGACCATGCGCGGAAACATGCTGAACAGATCCAGGAAGCCAGT 600

QY

1009

CGATTCCCAATCCGATTTGTGACGAGCAATAGCGTGAATACAAACAGCAACTCCCAACTA 1068

Db

601

CGATTCCCAATCCGATTTGTGACGAGCAATAGCTTGAATACAAACAGCAACTCCCAACTA 660

QY

1069

CGCGCGCGGAGGTGTGGCACCACAAAGTTCGACGACCACTGGATTACAGGTCACGGT 1128

```
Db 661 CGCCGGCGGAGTGTGGCACCACCAACGTCGACGACCACTGGATTACAGCTCCACGGT 720
QY 1129 ACTCGCAAGTGTCTACCGGCAGCAGCAGACNAGGATCG 1167
Db 721 ACTCGCAAGTGTCTACCGGCAGCAGCAGACNAGGATCG 759

RESULT 4
BF498461
LOCUS
DEFINITION
Drosophila melanogaster (fruit fly)
Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT12870 5 similar to CG12192:
FBan0012192 'motor protein' located on: 2R 59D1-59D1;: 04/09/2001,
mRNA sequence.
ACCESSION
BF498461
VERSION
BF498461.2 GI:13690358
KEYWORDS
Drosophila melanogaster (fruit fly)
SOURCE
Drosophila melanogaster
ORGANISM
Drosophila melanogaster
REFERENCE
1 (bases 1 to 763)
AUTHORS
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E.,
Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
Li,P., Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J.,
Pacleb,J., Pargass,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
Lewis,S.E., Celnikier,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
TITLE
BDGP/HMI AT Drosophila EST Project
JOURNAL
On Dec 6, 2000 this sequence version replaced gi:11581762.
COMMENT
Contact: Stapleton, M.

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003485: arm:X [10878159;1180508]
estimated-cyto:10A3-10B11: 04/09/2001
Plate: AT.128 row: F column: 10
High quality sequence stop: 673.

FEATURES
Location/Qualifiers
1..763
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT12870"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/notes="Organ: ADULT testes; Vector: pOTB7; Site 1: EcorI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

ORIGIN
Query Match 24.3%; Score 750.2; DB 2; Length 763;
Best Local Similarity 99.0%; Pred. No. 1e-183;
Matches 755; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1090 CCACACGTCGACGACCACTGGATTACAGGTCACCGTCTCCAGCTCTCCAGCTGTCTACCGGCC 1149
Db 1 CCACACGTCGACGACCACTGGATTACAGGTCACCGTCTCCAGCTGTCTACCGGCC 60

QY 1150 AGCAGCAGACCAAGATCGCTCGCGGTGCTTAATACACATTGCCCAATCCACGCGCG 1209
Db 61 AGCAGCAGACCAAGATCGCTCGCGGTGCTTAATACACATTGCCCAATCCACGCGCG 120
```

```
QY 1210 CAGCCAGTGTCTGTCCGGCGGCACAAGGAGTGCCTACCTCGGCCACAAACCCAGGAGCTG 1269
Db 121 CAGCCAGTGTCTGTCCGGCGGCACAAGGAGTGCCTACCTCGGCCACAAACCCAGGAGCTG 180
QY 1270 GCGCGCTAGTACCCCGCGCATCGCACGCAATTGAAAGAGGTGGAGCGGACTGAAGAGAAATC 1329
Db 181 GCGCGCTAGTACCCCGCGCATCTCACGCAATTGAAAGAGGTGGAGCGGACTGAAGAGAAATC 240
QY 1330 GCGAGAAGCGACGCGCCCGCACAGGCCGAGATGAAGAGAGAGAGTGGCGCTGATGAACC 1389
Db 241 GCGAGAAGCGACGCGCCCGCACAGGCCGAGATGAAGAGAGAGAGTGGCGCTGATGAACC 300
QY 1390 AGGATCCGCGCAATCCAAATCGGAGACGCGCGCAATGATACGCGAATATCAGAGCACGC 1449
Db 301 AGGATCCGCGCAATCCAAATCGGAGACGCGCGCAATGATACGCGAATATCAGAGCACGC 360
QY 1450 TGGAAATTTGTGCGCTGTCTCGATGGCGAGCGCGTTCGATGACCATCAGATCACAAGTGTGCG 1509
Db 361 TGGAAATTTGTGCGCTGTCTCGATGGCGAGCGCGTTCGATGACCATCAGATCACAAGTGTGCG 420
QY 1510 TGGCGAAGCGTCTCCATTAGCCGCAAGGAGGTCAATCGCAAGAGATCGATGTCAATTCGG 1569
Db 421 TGGCGAAGCGTCTCCATTAGCCGCAAGGAGGTCAATCGCAAGAGATCGATGTCAATTCGG 480
QY 1570 TGC CGCGCAAGGAGACATGCTCATCTGTCAGCGCGCGGAGGAGGTTCGACCTCACCAAGT 1629
Db 481 TGC CGCGCAAGGAGACATGCTCATCTGTCAGCGCGCGGAGGAGGTTCGACCTCACCAAGT 540
QY 1630 TCTGTGAGAACCAACAAGTTTCGTTCTGCACTAGCGCTTCAACGACACGTCGCGCAATGCCA 1689
Db 541 TCTGTGAGAACCAACAAGTTTCGTTCTGCACTAGCGCTTCAACGACACGTCGCGCAATGCCA 600
QY 1690 TGTATATCAAAATACACAGCAAGCGCTTGTGTAACCAATTTTCGAGGGCGGAATGCCGA 1749
Db 601 TGTATATCAAAATACACAGCAAGCGCTTGTGTAACCAATTTTCGAGGGCGGAATGCCGA 660
QY 1750 CGTGCTTCGCTTACCGCCAGAGCGGATCGGGCAAAACGACACACCATGGCGGTGAGTTTA 1809
Db 661 CGTGCTTCGCTTACCGCCAGAGCGGATCGGGCAAAACGACACCATGGCGGTGAGTTTA 720
QY 1810 ATGGAAGGTGCGAGGACTGCAAGAACCGCATCTACGCCATGGC 1852
Db 721 ATGGAAGGTGCGAGGACTGCAAGAACCGCATCTACGCCATGGC 763

RESULT 5
BF489668
LOCUS
DEFINITION
722 bp mRNA linear EST 23-APR-2001
AT25778.5prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT25778 5 similar to CSI453:
FBan0001453 'motor protein' located on: X 10A6-10A7;: 04/09/2001,
mRNA sequence.
ACCESSION
BF489668
VERSION
BF489668.2 GI:13755317
KEYWORDS
Drosophila melanogaster (fruit fly)
SOURCE
Drosophila melanogaster
ORGANISM
Drosophila melanogaster
REFERENCE
1 (bases 1 to 722)
AUTHORS
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E.,
Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
Li,P., Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J.,
Pacleb,J., Pargass,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
Lewis,S.E., Celnikier,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
TITLE
BDGP/HMI AT Drosophila EST Project
JOURNAL
On Dec 6, 2000 this sequence version replaced gi:11572969.
COMMENT
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
```

One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)  
 hit genomic ASO03483: arm:X [10878159..11180508]  
 estimated-cyto:10A3-10B11: 04/09/2001  
 Plate: AT.257 row: G column: 6  
 High quality sequence stop: 720.

## FEATURES

```

1. .722
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT25778"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT-10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/note="Organ: ADULT testes; Vector: pOTB7; Site 1: EGOR1;
Site 2: Xhol; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

```

## ORIGIN

Query Match	23.4%	Score 720.4	DB 2	Length 722
Best Local Similarity	99.9%	Pred. No. 5.9e-176		
Matches 721	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	423	AAATATTTTCGACCTAAAAAATCTGACTAAATTTGTTTGTGTTTATGTAATTTATTTA	482	
DB	1	AAATATTTTCGACCTAAAAAATCTGACTAAATTTGTTTGTGTTTATGTAATTTATTTA	60	
QY	483	GGCACATTTTGGCACACCAACGCTAGTGTACTACATCTACGACTTAACGGAACTCTCTCTGC	542	
DB	61	GGCACATTTTGGCACACCAACGCTAGTGTACTACATCTACGACTTAACGGAACTCTCTCTGC	120	
QY	543	AAGCAGTGGAAATTGCTGTCTCCATCAAGCAGTACTCGGAGTTAAACGACGAGTAAGCCGGGA	602	
DB	121	AAGCAGTGGAAATTGCTGTCTCCATCAAGCAGTACTCGGAGTTAAACGACGAGTAAGCCGGGA	180	
QY	603	GAAGAAGAAAGAGATCGGTGGAGATATAGATATACAGGTGGAGTCAAGAGGAAGATC	662	
DB	181	GAAGAAGAAAGAGATCGGTGGAGATATAGATATACAGGTGGAGTCAAGAGGAAGATC	240	
QY	663	ATGGACATGTTACGTTGGGCGAGCGCTCAAGATCAAGCGGACGATGCGCGGTCCAC	722	
DB	241	ATGGACATGTTACGTTGGGCGAGCGCTCAAGATCAAGCGGACGATGCGCGGTCCAC	300	
QY	723	ATGCCCGTGTGGCGGTGATCAACCACTGGGGAAGTGCAATCAAGTCGAATGGTACGAG	782	
DB	301	ATGCCCGTGTGGCGGTGATCAACCACTGGGGAAGTGCAATCAAGTCGAATGGTACGAG	360	
QY	783	CGCGGCGAAACGAAGGCGCAAGGAGGTAGAACTGGACGCCATCTCACTCAATCCGGAG	842	
DB	361	CGCGGCGAAACGAAGGCGCAAGGAGGTAGAACTGGACGCCATCTCACTCAATCCGGAG	420	
QY	843	CTAATCGAAGATACTGTCTGAAACGACGCGCGCCCGGAGCCGAAAGAACCGCCGC	902	
DB	421	CTAATCGAAGATACTGTCTGAAACGACGCGCGCCCGGAGCCGAAAGAACCGCTACCGCG	480	
QY	903	CCGATGAACCTCTCGCGTAAATCCACAATCGGCTATCGGTGGCAATCTCACACGCGGT	962	
DB	481	CCGATGAACCTCTCGCGTAAATCCACAATCGGCTATCGGTGGCAATCTCACACGCGGT	540	
QY	963	ATGACCATGCGCCGGAAACATGCTGAAACAAGATCCAGGAAGCCAGTTCGATTCGAATCCG	1022	
DB	541	ATGACCATGCGCCGGAAACATGCTGAAACAAGATCCAGGAAGCCAGTTCGATTCGAATCCG	600	
QY	1023	ATTGTCAGCAGCAATAGCGTGAATACAAACAGCAACTTCCAACACTACGCGCGCGGAGGT	1082	
DB	601	ATTGTCAGCAGCAATAGCGTGAATACAAACAGCAACTTCCAACACTACGCGCGCGGAGGT	660	

[illegible][illegible]

DR122830  
ACCESSION  
VERSION  
KEYWORDS  
EST.

SOURCE Drosophila pseudoobscura  
ORGANISM Drosophila pseudoobscura

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
1 (bases 1 to 1105)  
Richards S., Liu Y., Bettencourt B. R., Hradecky P., Ierovskiy S.

Nielsen, R., Thornton, K., Hubisz, M. J., Chen, R., Meisel, R. P., Couronne, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Bussemaker, H. J.,

van Batenburg, M. F., Howells, S. L., Scherer, S. E., Sodergren, E., Matthews, B. B., Crosby, M. A., Schroeder, A. J., Ortiz-Barrientos, D., and Ortíz-Barrientos, D. 2005. *Journal of Molecular Evolution* 61: 101-111.

Rives, C.M., Metzker, M.L., Muzny, D.M., Scott, G., Steffen, D., Wheeler, D.A., Worley, K.C., Havlak, P., Durbin, K.J., Egan, A.

Gill, R., Hume, J., Morgan, M. B., Miner, G., Hamilton, C., Huang, Y., Waldron, L., Verdusco, D., Clerc-Blankenburg, K. P., Dubchak, I.

Noor, M.A., Anderson, W., White, K.P., Clark, A.G., Schaeffer, S.W., Gelbart, W., Weinstock, G.M. and Gibbs, R.A.

**TITLE** Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution

JOURNAL  
PUBMED  
Genome Res. 15 (1), 1-18 (2005)  
15632085

**COMMENT**  
Contact: Stephen Richards  
Human Genome Sequencing Center

Baylor College of Medicine  
One Baylor Plaza, Houston, TX 77030, USA

**Tel: 713-798-6667**  
**Email: [stephen@bcm.tmc.edu](mailto:stephen@bcm.tmc.edu)**

NCBI Trace Archive: 226720539  
Insert Length: 1750 Std Error: 0.25.

```

FEATURES
  Location/Qualifiers
    source
      1. .ll05

```

```
/organism="Drosophila pseudoobscura"  
/mol_type="mRNA"
```

```

/db_xref="taxon:7237"
/clone="G9"

```

```
/dev stage="0-18h embryos"  
/clone lib="Drosophila pseudoobscura embryonic cDNA"
```

```
library"
/note="Vector: pDHE7: Site 1: EcoRI: Site 2: XhoI: oligo
```

ORIGIN  
dt priming from poly A+ RNA, directionally cloned<sup>a</sup>

Query Match	23.3%	Score 718	DB 8	Length 1105
-------------	-------	-----------	------	-------------

Best Local Similarity	83.0%	Pred. No. 2.8e-175;
Matches 853; Conservative	0;	Mismatches 172; Indels 3; Gaps 3;

QV 1466 GCTCGATGGCCAGGCCGTCGATGCCATCAGATCA CAG-TGTGGCTGGCGCAGCGCTCCA 1524

Db 2 GCACGAGGGTACGGCGGTGACGATCACCGCTTTGTGCGCAAGGTCAC 61

Qv 1525 TTAGCGCGCAAGGAGGTCATCGCAAGGAGATCGATGTCATTTCCGTGCCGCGCAAGGCA 1584

db 62 CAAGCCGAAAGGAGGTGAATCCCTAAGAGGTCGATGTCATATCGGTGCCCCCAAGGACT 121

Qv 1585 TGCTCATCGTGCACGAGCGCGCGACGAAGTTCGACCTCA CCAAGTTCCTGCGAACCA CA 1644

Db 122 TGATGATTGTCATAGCGCGGACCAAGGTGCGATCTCACAAAGTTCCTTGAACATCAC 181  
Qy 1645 AGTTTCCTTCGACTAGCCTTCAACACACGTGCGCAATGCATGATACAAATACA 1704  
Db 182 AATTTCGTTTCGATATGCTTCAACGAATGTGTCAATGCGATGCTGTACAAATACA 241  
Qy 1705 CAGCAAGCGTTGGTGAACACCATTTTCAGAGGCGGAATGGGACGTGCTTCGCCCTACG 1764  
Db 242 CAGCAAGCCCTGGTGAACACTATTTTCAGAGGCGGAATGGGACATGCTTTGCCCTATG 301  
Qy 1765 GCCAGACGGATCGGCAAAACGACACCATGCGCGGTGAGTTTAATGGAAGGTGCAGG 1824  
Db 302 GCCAGACGATCGGCAAGACCATACCATGCGCGGAGTTTACGGGTAAATACAGG 361  
Qy 1825 ACTGCAAGACGGATCTAGCCATGGCGCCCAAGGATGCTTTGTGACCTGAATATGC 1884  
Db 362 ACTGCAAGACGGATCTATGCCATGGCGCCCAAGGATGTTTCGGACCTCAATACAC 421  
Qy 1885 CGGTTTACCGCGCATGAATCTAGTCTCTCGGCCAGTTTCTTTTGAGATTTACAGTGGCA 1944  
Db 422 CGGCTATCGACCATGAATCTAGTCTCTCGGCCCTGTTCTTTGAGATTTACAGTGGCA 481  
Qy 1945 AGTCTTTCGATCTTCTGTCGACAGCAGAAACTCGCGCTCTCGAGGATGTTAAACAGC 2004  
Db 482 AGTCTTTCGATCTGTCGCCGATGAAGCAGAACTCGCGCTCTCGAGGACGGCAACAGC 541  
Qy 2005 AAGTGCAGGTGGGACTACCGAGAGGTGTCGATGCGCTCGAGGAGTACTGAGC 2064  
Db 542 AAGTGCAGGTGGTGGCTTCCAGGAAAGTTTGGACAGCGTCGAGGAGTCTAAAGA 601  
Qy 2065 TCATCAGCAGCGCAATGCTGCCGCAACATCCGCGCCAGAGCTCGGCCAACTCCAAATTCGT 2124  
Db 602 TCATCAGCAGCGCAATGCGCGCGCACCTCGGCGCCAGACCTCGGCCAATGCGAATCGT 661  
Qy 2125 CGCGTTTCGACGCGGCTTTCCAGATTTGCTGCGCGCCGACAGGCTCGACGAAGATCCATG 2184  
Db 662 CGCGATCGCATGCGCTCTCCAGATTTGCTGCGCTCGATGGGACGACAAAGATCCATG 721  
Qy 2185 GCAAGTTCGTTTCGATCTGCGCGGCAATGAGCGGCGCTGGAACATCTTCCTCGGCCG 2244  
Db 722 GTAAATTTTCGTTTCATGATCTGCGGGCAATGAGCGGGTGTGGACACCTCGTCGCGCG 781  
Qy 2245 ATCGCAGACGCTGATGAGGGTCCGAGATTAACAAATCGCTGCGGCCCTCAAGAGT 2304  
Db 782 ATGCCAGACCGATGAGGGGCGGAGATTAATGATCCCTGCTGGCACTTCAGGAGT 841  
Qy 2305 GCATTCGTTGGGCAACAGTCGGGCCCACTTGCCCTTCCTGCTCTCCAAACTCACCC 2364  
Db 842 GCATTCGGGCAATGGG-TAACAGTCGGGCCCACTGCGCTTCGTTCTCGAAGCTCAAGC 900  
Qy 2365 AGTGTCTGCGGACTGTTCAATGCGGAGAGAGAGAGTGCATGATAGCATGATCT 2424  
Db 901 ANGTGTGCGGACTGTTCAATGGAGAGAGCAAGACTTGCATGATAACCATGATAT 960  
Qy 2425 CGCGGGACTTACTCTCTGAGCAGACGCTCAACGCTGCGCTATGCGGATCGTGTCA 2484  
Db 961 CGCGGCTG-CAACATTTCTGTGAGCATACGCTCAACAGCTGCGCTTACGGAATCGGGCT 1019  
Qy 2485 AGGAGCTG 2492  
Db 1020 AGGAACGTG 1027

RESULT 7  
DR135726 1153 bp mRNA linear EST 16-JUN-2005  
LOCUS 49291254 Drosophila pseudoobscura embryonic cDNA library Drosophila  
DEFINITION pseudoobscura cDNA clone I3 5', mRNA sequence.  
ACCESSION DR135726  
VERSION DR135726.1 GI:67880816  
KEYWORDS EST.  
SOURCE Drosophila pseudoobscura

## ORGANISM

Drosophila pseudoobscura  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

1 (bases 1 to 1153)

## AUTHORS

Richards,S., Liu,Y., Bettencourt,B.R., Hradecky,P., Letovsky,S.,  
Nielsen,R., Thornton,K., Hubisz,M.J., Chen,R., Meisel,R.P.,  
Couronne,O., Hua,S., Smith,M.A., Zhang,P., Liu,J., Bussemaker,H.J.,  
van Batenburg,M.F., Howells,S.L., Scherer,S.B., Sodergren,E.,  
Matthews,B.B., Crosby,M.A., Schroeder,A.J., Ortiz-Barrientos,D.,  
Rives,C.M., Metzker,M.L., Muzny,D.M., Scott,G., Steffen,D.,  
Wheeler,D.A., Worley,K.C., Havlak,P., Durbin,K.J., Egan,A.,  
Gill,R., Hume,J., Morgan,M.B., Miner,G., Hamilton,C., Huang,Y.,  
Waldron,L., Verdusco,D., Clerc-Blankenburg,K.P., Dubchak,I.,  
Noor,M.A., Anderson,W., White,K.P., Clark,A.G., Schaeffer,S.W.,  
Gelbart,W., Weinscock,G.M. and Gibbs,R.A.  
Comparative genome sequencing of Drosophila pseudoobscura:  
chromosomal, gene, and cis-element evolution  
Genome Res. 15 (1), 1-18 (2005)

## TITLE

15632085

## JOURNAL

15632085

## COMMENT

Contact: Stephen Richards  
Human Genome Sequencing Center  
Baylor College of Medicine  
One Baylor Plaza, Houston, TX 77030, USA  
Tel: 713-798-6667  
Email: stephenr@bcm.tmc.edu  
NCBI Trace Archive: 226715462  
Insert Length: 1750 Std Error: 0.25.

## FEATURES

## source

1..1153  
/organism="Drosophila pseudoobscura"  
/mol\_type="mRNA"  
/db\_xref="taxon:7237"  
/clone="I3"  
/dev\_stage="0-18h embryos"  
/clone\_lib="Drosophila pseudoobscura embryonic cDNA  
library"  
/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo  
dt priming from poly A+ RNA, directionally cloned"

## ORIGIN

Query Match 22.5%; Score 692; DB 8; Length 1153;  
Best Local Similarity 82.9%; Pred. No. 1.7e-168;  
Matches 813; Conservative 0; Mismatches 166; Indels 2; Gaps 2;  
Qy 1452 GAATTTGTCGCGTGTGATGCGCAGGCGGTGAGATCAATCAGATCA-CAGTGTGCGT 1510  
Db 19 GACCTGGTGNATCGGCACGAGGTCGCGGCTTGACGATCACGAGATCACCCCTTTATGT 78  
Qy 1511 GCGCAAGCGTCCCATTTAGCGCAAGGAGTCAATCGCAAGGAGATCGATGTCATTTCCGT 1570  
Db 79 GCGCAAGCGTCCCAAGCGCAAGGAGTGAATCTTTAGAGGTGATGTATATCGGT 138  
Qy 1571 GCGCGCAAGGACATCTCTCATCGTGCAAGCGCGCGCAGCAGGTCGACCTCACCAAGTT 1630  
Db 139 GCGCGCAAGGACTTGATGATTTGTCATGAGCGCGGACCAAGTCTGATCTCACCAAGTT 198  
Qy 1631 CTTGGAGAACCAAGTTTTCGTTTCGATCTGACCTTCAACGACATGCGGACAAATGCCAT 1690  
Db 199 CTTGAAACATCAAAATTTTCGTTTCGACTATGCTTCAACGAAATGTGTGACAATGCCAT 258  
Qy 1691 GGTATACAAATACAGCGCAAGCGGTTGTTGTTGAAAACCAATTTTCGAGGCGGAATGGCAC 1750  
Db 259 GGTGTACAAATACAGCGCAAGCGGCTGTTGTTGAAAACCAATTTTCGAGGCGGAATGGCAC 318  
Qy 1751 GTGCTTCGCTACCGGCACAGCGGATCGGGCAAAACGACACCATCGGCGGTGAGTTTAA 1810  
Db 319 ATGCTTTGCTATGGCAGACAGGATCCGCGCAAGACCCATACCATGCGGGGAGTTTAC 378  
Qy 1811 TGGAAAGGTGCGGACTGCAAGAACGGCATCTACGCCATGCGGCGCAAGATGCTTTTGT 1870  
Db 379 GGGTAAATACAGGACTGCAAGAACGGCATCTATGCCATGCGGCGCAAGATGCTTTCG 438







```
FEATURES
  source
    High quality sequence stop: 643.
    Location/Qualifiers
      1..687
        /organism="Drosophila melanogaster"
        /mol_type="mRNA"
        /db_xref="taxon:7227"
        /clone="AT04507"
        /sex="male"
        /dev_stage="0-3 day old Ore-R males"
        /lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
          AT.121-AT.319: DH5-alpha Tona"
        /clone_lib="AT Drosophila melanogaster adult testes pOTB7"
        /note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
          Site_2: XhoI; The mRNA for the testis library was made
          from testes and seminal vesicles hand dissected from 0-3
          day old Ore-R males. RNA kindly provided by the lab of
          Margaret Fuller. Sized fractionated cDNAs were directly
          ligated into pOTB7. Plasmid cDNA library."
ORIGIN
  Query Match      22.1%; Score 680.6; DB 2; Length 687;
  Best Local Similarity 99.4%; Pred. No. 1.3e-165;
  Mismatches 683; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
  QY 388 CACAAGTTTATCGCTCGTCAAGAACACACGAAATAAAATTTATTTTCGACCTAAAAATCT 447
  Db 1 CACAAGTTTATCGCTCGTCAAGAACACACGAAATAAAATTTATTTTCGACCTAAAAATCT 60
  QY 448 GACTAAATTTGTTTTTTTGTATTTATTTATTTAGGCACATTTTGACACACCAACGTA 507
  Db 61 GACTAAATTTGTTTTTTTGTATTTATTTATTTAGGCACATTTTGACACACCAACGTA 120
  QY 508 GTTACTACATCTACGACTAACGGAATCTCTCTCGAAGCAGTGGAAGTTGTCCTCATCA 567
  Db 121 GTTACTACATCTACGACTAACGGAATCTCTCTCGAAGCAGTGGAAGTTGTCCTCATCA 180
  QY 568 AGCAGTACTCGAGTTAAGCAGCATTAAGCCGGAGAAAGAGAGATCGGTGGAGAA 627
  Db 181 AGCAGTACTCGAGTTAAGCAGCATTAAGCCGGAGAAAGAGAGATCGGTGGAGAA 240
  QY 628 TAGAGATATACAGTGGAGTCAAGAGAAAGGATCATGGACATGATTACCGTGGGGCAGA 687
  Db 241 TAGAGATATACAGTGGAGTCAAGAGAAAGGATCATGGACATGATTACCGTGGGGCAGA 300
  QY 688 CGGTCAAGATCAAGCGGACGATGGCGGTGTCATGGCGGTGATGATCAACC 747
  Db 301 CGGTCAAGATCAAGCGGACGATGGCGGTGTCATGGCGGTGATGATCAACC 360
  QY 748 AGTCGGCAAGTGCATCACAGTGAATGTTACGAGCGCGGGAACCAAGGGCAAGGAG 807
  Db 361 AGTCGGCAAGTGCATCACAGTGAATGTTACGAGCGCGGGAACCAAGGGCAAGGAG 420
  QY 808 TAGAATCGGACGCGCATCTACGCTCAATCGGAGCTTAATGCAAGATCTGTGCAACAGC 867
  Db 421 TAGAATCGGACGCGCATCTACGCTCAATCGGAGCTTAATGCAAGATCTGTGCAACAGC 480
  QY 868 ACGCGCGCGCGGAGCCCAAGAAACAAGCCACCGCGCGGATGAACCTTCGCGTAATCCCA 927
  Db 481 ACGCGCGCGCGGAGCCCAAGAAACAAGCCACCGCGCGGATGAACCTTCGCGTAATCCCA 540
  QY 928 CACAATCGGCTATCGGTGGCAATCTCACAGCCGCTATGACATGGCCGGAACATGCTGA 987
  Db 541 CACAATCGGCTATCGGTGGCAATCTCACAGCCGCTATGACATGGCCGGAACATGCTGA 600
  QY 988 ACAAGATCCAGGAAGCCAGTCCGATTCCTCAATCCGATTGTGACAGCAATAGCGTGAATA 1047
  Db 601 ACAAGATCCAGGAAGCCAGTCCGATTCCTCAATCCGATTGTGACAGCAATAGCGTGAATA 660
  QY 1048 CAAACAGCAATCTCAACACTACGCGCG 1074
  Db 661 CAAACAGCAATCTCAACACTACGCGCG 687
```

```
RESULT 11
BF499204
LOCUS
DEFINITION
  BF499204      886 bp      mRNA      linear      EST 19-APR-2001
  AT13721.5prime AT Drosophila melanogaster adult testes pOTB7
  Drosophila melanogaster cDNA clone AT13721.5 similar to CG1453;
  Fban0001453 'motor protein' located on: X 10A6-10A7; 04/09/2001,
  mRNA sequence.
ACCESSION
  BF499204
VERSION
  BF499204.2 GI:13691069
KEYWORDS
  EST.
SOURCE
  Drosophila melanogaster (fruit fly)
  Drosophila melanogaster
  Sukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 886)
  Stapleton,M., Brokstein,P., Hong,L., Agbavani,A., Baxter,E.,
  Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
  Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
  Li,P., Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J.,
  Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
  Lewis,S.E., Celnikier,S. and Rubin,G.M.
  BDGP/HMI AT Drosophila EST Project
  Unpublished (2000)
  On Dec 6, 2000 this sequence version replaced gi:11582505.
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AEO03485: arm:X [10878159..11180508]
  estimated-cyto:10A3-10B11: 04/09/2001
  Plate: AT.137 row: B column: 9
  High quality sequence stop: 609.
FEATURES
  source
    Location/Qualifiers
      1..686
        /organism="Drosophila melanogaster"
        /mol_type="mRNA"
        /db_xref="taxon:7227"
        /clone="AT13721"
        /sex="male"
        /dev_stage="0-3 day old Ore-R males"
        /lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
          AT.121-AT.319: DH5-alpha Tona"
        /clone_lib="AT Drosophila melanogaster adult testes pOTB7"
        /note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
          Site_2: XhoI; The mRNA for the testis library was made
          from testes and seminal vesicles hand dissected from 0-3
          day old Ore-R males. RNA kindly provided by the lab of
          Margaret Fuller. Sized fractionated cDNAs were directly
          ligated into pOTB7. Plasmid cDNA library."
ORIGIN
  Query Match      22.0%; Score 678; DB 2; Length 686;
  Best Local Similarity 99.3%; Pred. No. 6.4e-165;
  Mismatches 681; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
  QY 388 CACAAGTTTATCGCTCGTCAAGAACACACGAAATAAAATTTATTTTCGACCTAAAAATCT 447
  Db 1 CACAAGTTTATCGCTCGTCAAGAACACACGAAATAAAATTTATTTTCGACCTAAAAATCT 60
  QY 448 GACTAAATTTGTTTTTTTGTATTTATTTATTTAGGCACATTTTGACACACCAACGTA 507
  Db 61 GACTAAATTTGTTTTTTTGTATTTATTTATTTAGGCACATTTTGACACACCAACGTA 120
  QY 508 GTTACTACATCTACGACTAACGGAATCTCTCTCGAAGCAGTGGAAGTTGTCCTCATCA 567
  Db 121 GTTACTACATCTACGACTAACGGAATCTCTCTCGAAGCAGTGGAAGTTGTCCTCATCA 180
  QY 568 AGCAGTACTCGAGTTAAGCAGCATTAAGCCGGAGAAAGAGAGATCGGTGGAGAA 627
  Db 181 AGCAGTACTCGAGTTAAGCAGCATTAAGCCGGAGAAAGAGAGATCGGTGGAGAA 240
```

/note="vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
fractionated cDNAs were directly ligated into pOT2.  
Plasmid cDNA library."

241	Db	TAGAGATATACAGGTGGAGTCAAGAGGAGGAGCATGGACATGATTACGGTGGGGCAGA	300
688	Qy	CGGTCAAGATCAACGGGACGGATGGCCGCGTCCACATGGCCGTTGGTGGCGGTGATCAACC	747
301	Db	CGGTCAAGATCAACGGGACGGATGGCCGCGTCCACATGGCCGTTGGTGGCGGTGATCAACC	360
748	Qy	AGTCGGGCAGTGTGATCACAGTCCAAATGGTACGAGCGGGCGGAAACGAAAGGGCAAGGAGG	807
361	Db	AGTCGGGCAGTGTGATCACAGTCCAAATGGTACGAGCGGGCGGAAACGAAAGGGCAAGGAGG	420
808	Qy	TAGAACTGGAGCCCATCTACACGTCTCAATCCGGAGCTAAATCCAAGATACTCTTCGAAACAGC	867
421	Db	TAGAACTGGAGCCCATCTACACGTCTCAATCCGGAGCTAAATCCAAGATACTCTTCGAAACAGC	480
868	Qy	ACGCGCGCCCGGAGGCCCAAGAAAACAAGCCACCGCGCCGATGAACCTCTCGCGTAAATCCCA	927
481	Db	ACGCGCGCCCGGAGGCCCAAGAAAACAAGCCACCGCGCCGATGAACCTCTCGCGTAAATCCCA	540
928	Qy	CACAATCGGCTATCGGTGGCAATCTCACACGCGCTATGACCATGGCGGGAACATGCTGA	987
541	Db	CACAATCGGCTATCGGTGGCAATCTCACACGCGCTATGACCATGGCGGGAACATGCTGA	600
988	Qy	ACAAGATCCAGGAAAGCCAGTCGATTCCCAATCCGATTGTCAGCAGCAATAGCCGTGAATA	1047
601	Db	ACAAATCCAGGAAAGCCAGTCGATTCCCAATCCGATTGTCAGCAGCAATAGCTTGAATA	660
1048	Qy	CAAAACAGCAACTCCAAACACTACGGCC	1073
661	Db	CAAAACAGCAACTCCAAACACTACGGGC	686

RESULT 12					
B163599					
LOCUS	B163599	689 bp	mRNA	linear	EST 10-SEP-2001
DEFINITION	Sp16989.Sprime SD Drosophila melanogaster Schneider L2 cell culture port2 Drosophila melanogaster cDNA clone Sp16989 5 similar to Kjp10A: FBAN001453 GO:[kinesin (GO:0005871); motor protein (GO:000374)] located on: X 10A6-10A7:: 05/18/2001, mRNA sequence.				

EST.  
 Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 689)  
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
 Lewis, S. and Rubin, G.M.  
 BGP/HMI Drosophila EST Project  
 Unpublished (2001)  
 Contact: Stapleton, M.  
 BDGP

One Lyncrotron RD, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 hit genomic AE003485: arm:X [10878159,11180508]  
 estimated-cyto:10A3-10B11: 05/18/2001  
 Plate: SD.169 row: H column: 5  
 High quality sequence stop: 679.  
 Location/Qualifiers  
     1. .689  
     /organism="Drosophila melanogaster"  
     /mol\_type="mRNA"  
     /db\_xref="taxon:7227"  
     /clone="SD16899"  
     /lab\_host="DH5-alpha"  
     /clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
     culture pot2"

```
REFERENCE
AUTHORS
1 (bases 1 to 679)
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E.,
Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
Farfan,D., Friese,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
Li,P., Liao,G., Miranda,A., Miera,S., Mungall,C.J., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
Lewis,S.E., Celniker,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
JOURNAL
COMMENT
On Dec 6, 2000 this sequence version replaced gi:11578538.
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003485: arm:X [10878159,11180508]
estimated-cyto:10A3-10B11: 04/07/2001
Plate: AT.41 row: B column: 2
High quality sequence stop: 655.
FEATURES
Location/Qualifiers
1..679
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT04114"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes potB7"
/notes="Organ: ADULT testes; Vector: potB7; Site 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into potB7. Plasmid cDNA library."
ORIGIN
Query Match 21.5%; Score 663.6; DB 2; Length 679;
Best Local Similarity 99.0%; Pred. No. 3.2e-161;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 388 CACAAGTTTATCGTCTGCAAGAAACAAACGAAATAAAATTTATTTTCGACCTAAAAATCT 447
DB 1 CACAAGTTTATCGTCTGCAAGAAACAAACGAAATAAAATTTATTTTCGACCTAAAAATCT 60
QY 448 GACTAATATGTTTGTGTTTATGTTATTTATTTAGGCACATTTTGCACACCAACGTA 507
DB 61 GACTAATATGTTTGTGTTTATGTTATTTATTTAGGCACATTTTGCACACCAACGTA 120
QY 508 GTTACTATCATCTACGACTAACGGAACTCTCTCTCAAGCAGTGGAACTGCTGTCCATCA 567
DB 121 GTTACTATCATCTACGACTAACGGAACTCTCTCTCAAGCAGTGGAACTGCTGTCCATCA 180
QY 568 AGCAGTACTCGGAGTTAACCCAGGATAAGCCGGGAGAAAGAGATCGGTGGAGAA 627
DB 181 AGCAGTACACGGAGTTAACCCAGGATAAGCCGGGAGAAAGAGATCGGTGGAGAA 240
QY 628 TAGAGATATACAGTGGAGTCAAGAGGAGGATCATGGACATGATTACGTTGGGCGAGA 687
DB 241 TAGAGATATACAGTGGAGTCAAGAGGAGGATCATGGACATGATTACGTTGGGCGAGA 300
QY 688 CGCTCAAGATCAAGCGGACGGATGGCCCGCTGCACATGGCCGCTGGTGGCGGTGATCAACC 747
DB 301 CGCTCAAGATCAAGCGGACGGATGGCCCGCTGCACATGGCCGCTGGTGGCGGTGATCAACC 360
QY 748 AGTCGGGCAAGTGCATCACAGTGCAGTGGATGGTACGAGCGCGCGAAACGAAAGGCAAGGAGG 807
DB 361 AGTCGGGCAAGTGCATCACAGTGCAGTGGATGGTACGAGCGCGCGAAACGAAAGGCAAGGAGG 420
QY 808 TAGAATCGGACGCCATCTACCGTCAATCGGAGCTAATGCAAGATCTGTGCAACAGC 867
Db 421 TAGAATCGGACGCCATCTACCGTCAATCGGAGCTAATGCAAGATCTGTGCAACAGC 480
QY 868 ACGCCGCCCGCGAGCCCAAGAAACAAAGCACCGCGCGATGAACCTCTCGCGTAAATCCCA 927
Db 481 ACGCCGCCCGCGAGCCCAAGAAACAAAGCACCGCGCGATGAACCTCTCGCGTAAATCCCA 540
QY 928 CACAATCGGCTATCGGTGGCAATCTCACCGACCGTATGACCATCGCCGGAACAATGCTGA 987
Db 541 CACAATCGGCTATCGGTGGCAATCTCACCGACCGTATGACCATCGCCGGAACAATGCTGA 600
QY 988 ACAGATGCCAGGAAGACCAAGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 1047
Db 601 ACAAATCCAGGAAGACCAAGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 660
QY 1048 CAAACAGCAACTCCA 1062
Db 661 ACAACAGCAACTCCA 675
RESULT 14
LOCUS
BG634444 670 bp mRNA linear EST 23-APR-2001
DEFINITION
AT30655.5prime AT Drosophila melanogaster adult testes potB7
Drosophila melanogaster cDNA clone AT30655 5 similar to CG1453:
FBan0001453 'motor protein' located on: X 10A6-10A7; 04/09/2001,
mRNA sequence.
ACCESSION
BG634444
VERSION
BG634444.1 GI:13761981
KEYWORDS
EST.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 670)
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E.,
Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
Farfan,D., Friese,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
Li,P., Liao,G., Miranda,A., Miera,S., Mungall,C.J., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
Lewis,S.E., Celniker,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003485: arm:X [10878159,11180508]
estimated-cyto:10A3-10B11: 04/09/2001
Plate: AT.306 row: E column: 7
High quality sequence stop: 615.
FEATURES
Location/Qualifiers
1..670
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT30655"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes potB7"
/notes="Organ: ADULT testes; Vector: potB7; Site 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into potB7. Plasmid cDNA library."
ORIGIN
Query Match 21.5%; Score 663.6; DB 2; Length 670;
```

Best Local Similarity 99.4%; Pred. No. 3.6e-161; Matches 666; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 379 CTTAGTTTCACAAAGTTTATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCACCT 438

Db 1 CTTAGTTTCACAAAGTTTATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCACCT 60

QY 439 AAAAAATCTGACTAAATGTGTTTTTTTGTATGTATTTATTTAGGCACATTTTGACAC 498

Db 61 AAAAAATCTGACTAAATGTGTTTTTTTGTATGTATTTATTTAGGCACATTTTGACAC 120

QY 499 CACAACGTAGTTACTACATCTACGACTTAACGGAATCTCTCTCAAGACGATGGAAAGTTGC 558

Db 121 CACAACGTAGTTACTACATCTACGACTTAACGGAATCTCTCTCAAGACGATGGAAAGTTGC 180

QY 559 TGTCCATCAACGAGTACTCGAGTTAAACGAGATAAGCCGGGAGAAAGAAAGATC 618

Db 181 TGTCCATCAACGAGTACTCGAGTTAAACGAGATAAGCCGGGAGAAAGAAAGATC 240

QY 619 GGTGGAGATAGAGATATACAGGTGGAGTCAAGAGAGAAAGATCATGGACATGTTACGS 678

Db 241 GGTGGAGATAGAGATATACAGGTGGAGTCAAGAGAGAAAGATCATGGACATGTTACGS 300

QY 679 TGGGGCAGACGCTCAAGATCAACGCGACGATGCGCGCTCCACATGCGCGTGGTGGCGG 738

Db 301 TGGGGCAGACGCTCAAGATCAACGCGACGATGCGCGCTCCACATGCGCGTGGTGGCGG 360

QY 739 TGATCAACCAAGTCGGGCAAGTGCATCACAGTCAAGTGGTACAGAGCCGGCGAAACGAAG 798

Db 361 TGATCAACCAAGTCGGGCAAGTGCATCACAGTCAAGTGGTACAGAGCCGGCGAAACGAAG 420

QY 799 GCAAGGAGGTAGAACTGGAGCCCATCTACGCTCAATCGGAGCTTAATGCAAGATACG 858

Db 421 GCAAGGAGGTAGAACTGGAGCCCATCTACGCTCAATCGGAGCTTAATGCAAGATACG 480

QY 859 TCGAACACGACGCGCGCCCGGAGCCCAAGAAACAAGCCACCGCGCGGATGAACCTCTCGC 918

Db 481 TCGAACACGACGCGCGCCCGGAGCCCAAGAAACAAGCCACCGCGCGGATGAACCTCTCGC 540

QY 919 GTAATCCACACAATCGGCTATCGGTGGCATCTCAGCAGCGTATGACATGGCCGGAA 978

Db 541 GTAATCCACACAATCGGCTATCGGTGGCAATCTCAGCAGCGTATGACATGGCCGGAA 600

QY 979 ACATGCTGAACAGATCCAGGAAGCCAGTCGATTCCCAATCCGATTGTTCAGCAGCAATA 1038

Db 601 ACATGCTGAACAGATCCAGGAAGCCAGTCGATTCCCAATCCGATTGTTCAGCAGCAATA 660

QY 1039 CGCTGAATAC 1048

Db 661 GCTTGAATAC 670

RESULT 15

BI609735

LOCUS

DEFINITION

BI609735 664 bp mRNA linear EST 07-SEP-2001

RH14556.Sprime RH Drosophila melanogaster normalized head pFlc-1

Drosophila melanogaster cDNA clone RH14556 5 similar to Klp10A;

PBano001453 GO:[kinesin (GO:0005871)]; motor protein (GO:0003774)]

located on: X 10A6-10A7; 08/17/2001, mRNA sequence.

BI609735

BI609735.1 GI:15505260

EST.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Phylloidea; Drosophilidae; Drosophila.

1 (bases 1 to 664)

Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and

Rubin,G.M.

BDGP/RHMI RH Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu

hit genomic AB003485; arm:X [10878159,11180508]

estimated-cyto:10A3-10B11; 08/17/2001

Plate: RH.145 row: E column: 8

High quality sequence stop: 563.

Location/Qualifiers

1..564

/organism="Drosophila melanogaster"

/mol\_type="mRNA"

/db\_xref="taxon:7227"

/clone="RH14556"

/sex="male and female"

/dev\_stage="Adult"

/lab\_host="DHS-alpha TonA"

/clone\_lib="RH Drosophila melanogaster normalized head pFlc-1"

/note="Organ: head; Vector: pFlc1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

ORIGIN

Query Match 21.5%; Score 662.4; DB 3; Length 664;

Best Local Similarity 99.8%; Pred. No. 7.4e-161;

Matches 663; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 506 TAGTTACTACATCTACGACTAACGGAATCTCTCTGCAAGCAGTGGAGTTGCTGTCCAT 565

Db 1 TAGTTACTACATCTACGACTAACGGAATCTCTCTGCAAGCAGTGGAGTTGCTGTCCAT 60

QY 566 CAAGCAGTACTCGGAGTTAAACGAGATAAGCCGGGAGAAAGAGATCGGTGGAG 625

Db 61 CAAGCAGTACTCGGAGTTAAACGAGATAAGCCGGGAGAAAGAGATCGGTGGAG 120

QY 626 AATAGAGATATACAGTGGAGTCAAGAGAGAGATCATGGACATGATTACGGTGGGCA 685

Db 121 AATAGAGATATACAGTGGAGTCAAGAGAGAGATCATGGACATGATTACGGTGGGCA 180

QY 686 GAGCGTCAAGATCAAGCGGACGATGGCGCGTCCACATGGCCGTGGTGGCGTGAATCA 745

Db 181 GAGCGTCAAGATCAAGCGGACGATGGCGCGTCCACATGGCCGTGGTGGCGTGAATCA 240

QY 746 CCAGTCCGGGCAAGTGCATCAGTCCGATGGTACGAGCGCGGCAAGAGGCGCAAGGA 805

Db 241 CCAGTCCGGGCAAGTGCATCAGTCCGATGGTACGAGCGCGGCAAGAGGCGCAAGGA 300

QY 806 GGTAGAACTGGAAGCCATACCTCACCTCAATCCGAGCTAATGCAAGATCATGTGCGAACA 865

Db 301 GGTAGAACTGGAAGCCATACCTCACCTCAATCCGAGCTAATGCAAGATCATGTGCGAACA 360

QY 866 GCACCGCCCGCCGGAGCCCAAGAAACAAGCCACCGCGCGATGAAACCTCTCGCGTAATCC 925

Db 361 GCACCGCCCGCCGGAGCCCAAGAAACAAGCCACCGCGCGATGAAACCTCTCGCGTAATCC 420

QY 926 CACACAATCGGTATCGGTGGCAATCTCACCGCCGTATGACCATGGCCGGGAAACATGCT 985

Db 421 CACACAATCGGTATCGGTGGCAATCTCACCGCCGTATGACCATGGCCGGGAAACATGCT 480

QY 986 GAAACAAGATCCAGGAAGCCAGTCCGATTCCCAATCCGATTGTTCAGCAGCAATAGCGTAA 1045

Db 481 GAAACAAGATCCAGGAAGCCAGTCCGATTCCCAATCCGATTGTTCAGCAGCAATAGCGTAA 540

QY 1046 TACAACACGCAACTCCAAACATCTACCGCGCGGAGGAGTGGTGGCAACCAACCTTCGACGAC 1105

Db 541 TACAACACGCAACTCCAAACATCTACCGCGCGGAGTGGTGGCAACCAACCTTCGACGAC 600

Qy	1106	CACTGGATTACAGCGTCCACGGTACTCGCAAGCTGCTACCGGCCAGCAGCAGACAAGCAT	1165
Db	601	CACTGGATTACAGCGTCCACGGTACTCGCAAGCTGCTACCGGCCAGCAGCAGACAAGCAT	660
Qy	1166	CGCC	1169
Db	661	CGCC	664

Search completed: May 12, 2006, 21:09:21  
 Job time : 11794 secs

***This Page Blank (uspto)***